

When will cigarettes be smoked out?

The restriction of tobacco smoking in the United States, already severe, is likely soon to be taken further and, in the process, to demonstrate that tobacco and cannabis cannot be dealt with differently.

THE tobacco industry is having a torrid time, at least in the United States. It is not just that the apparently inexorable spread of the familiar "No Smoking!" signs has now reached most buildings occupied by people unrelated to each other, but that the industry appears to be losing the support it used to be able to command in the US Congress. So much should be clear from last week's events (see page 696).

What has happened is that the weight of public opinion hostile to tobacco is now such that the administration suspects it could win support for draconian regulation; powerful officials are asking ruder questions than even a few years ago. The Congress seems happy to collaborate. Last week, it staged an intriguing dispute between the Food and Drug Administration and the industry about the need for double-strength tobacco, measured by nicotine content.

In themselves, these events are not all that remarkable; tobacco manufacturers have been increasingly frequent witnesses at congressional committees in the past few years. The novelty is that they now have so few friends. Much less is now heard, for example, of the economic case for tobacco — that it provides employment in otherwise rural parts and that it keeps tens of thousands of corner shops and kiosks in business. At the same time, the industry's enemies have grown in number and ferocity. And what they are now saying is likely to be damaging to the industry. Even if double-strength cigarettes have not been widely sold, for example, their mere existence will seem a folly to those concerned with the public health (not to mention their own). And the opposition to tobacco is going to get stronger.

Language

In the circumstances, it is curious that the industry has not learned to adapt its language to the changed circumstances. It now carries little weight to say (as one spokesman did last week) that smoking is only one of several "risk factors" for mortality, alongside diet and lifestyle. Or to insist that nicotine is not addictive, but to allow that cigarette-smoking may be "habit-forming". The plain truth is that ordinary people, including those who smoke cigarettes, know that smoking predisposes people towards lung cancer, respiratory and cardiovascular trouble and (with lower correlation) towards cancer more generally. And would those hooked by cigarettes spend all that money on hypnotists and other supposed remedies for smoking if their habits were easily distinguishable from addiction?

The difficulty, for the tobacco industry, is that what it says

in its own defence is constrained less by reality than by the perceived dangers of being more open. The fine distinction between habit and addiction, for example, is drawn because, if tobacco were acknowledged to be addictive, it could more readily be classified as a drug akin to cannabis, tobacco tycoons would be open to arrest by the US Drug Enforcement Agency. Nor can the industry admit that the use of tobacco may harm people's health without increasing the likelihood that the long queue of pending lawsuits brought on behalf of smokers or their surviving relatives will turn into a flood of successful personal damage awards by the courts. And even to claim what is the truth about the tobacco industry, that the continued sale of cigarettes is justified by people's wish to buy them, would evoke irresistible demands for a ban on advertisements for tobacco. The industry is in a cleft stick. What else can it expect?

Interdiction

But so is the US government and its counterparts elsewhere. Nowhere (unless in some Moslem country) in the near future will there be a strict legal interdiction of the sale of tobacco. Tobacco growers would indeed be harmed and other people put out of work. Worse, smugglers would prosper, perhaps so much that they compared in power and influence with the handlers of cannabis, cocaine and heroin. But such a law would be a gross interference with people's liberty to do what they wish with their personal lives. Governments accept the first two arguments, but not the third, because it applies to cannabis as well as to tobacco. But that shows that the two substances should be dealt with on an equal footing.

What should that be? If tobacco (cannabis) is harmful, people should be told about the dangers. So, as with AIDS, for example, governments should advertise on their own account, on television and elsewhere, to warn their electors of the risks about them. If people hanker so after tobacco (cannabis) that they will pay large sums of money to buy it illegally, governments should be prepared to issue tobacco (cannabis) permits for the purchase of limited supplies at state-run shops. (That is how alcohol must be bought in many US states.) Permits would of course be renewed only if applicants produced medical or other evidence that they had not yet been harmed by their use of a licensed substance. Commercial advertising of these materials would of course be banned. That is where governments should be aiming with tobacco (cannabis). The tobacco industry, which has only the most meagre social case to make for continued



existence, should also learn to live with that prospect.

When that state of affairs will come about, and what it will be like, remains anybody's guess. Despite the recent unexpected growth in the United States of the opposition to tobacco, the difficulties of carrying contentious issues such as this into law are huge. It is not just that the manufacturers are powerful corporations, but that the difficulty in striking a proper balance between public health and civil liberty are themselves substantial. But tobacco and cannabis create problems that differ only in their origin and degree. Sooner or later they will have to be dealt with together. Before then, the tobacco companies will almost certainly be shrinking under pressure from the courts. □

Palaeontology cleanup

The University of Chandigarh now has a chance to shed the dubious reputation in palaeontology its has acquired.

It is a relief, but in some ways also a surprise, that the committee of inquiry at Chandigarh appears to have found that Professor V. J. Gupta, the palaeontologist, did indeed fabricate many of the data on which his many publications have been based (see page 698). Gupta has always stoutly defended himself against criticism, but fears that his and his friends' influence would lead the inquiry in a different direction were clearly not justified. That is the good news. The other side of the coin is that the judicial inquiry into Gupta's publications has taken two years to establish what was widely accepted in 1989. How long will it take the University of Chandigarh to decide what should happen to Gupta and his post in the geology department? And what, for that matter, should it decide?

What stands out is that the past five years have been a tragedy for the Panjab University, at which Gupta is a senior professor. The university has lost respect in the outside world, and has suffered financially. But people have also suffered. Gupta himself, despite the lifting of his suspension from his post, has not been taken seriously for the past five years and is unlikely, now, to win back his reputation. But the victims most deserving of sympathy are Gupta's less senior co-workers at Chandigarh, three of whom were also subject to the inquiry carried out by M. S. Gujral (but have been found free from taint). Their research has been interrupted just as Gupta's, but they have also had to put up with accusations of disloyalty and worse.

As a consequence, the University of Chandigarh cannot now be exclusively concerned with its own future reputation; there is also a matter of personal justice to decide. But these considerations point in the same direction, and to the conclusion that Gupta should leave the university. That is a hard thing to say of one whose whole academic life has been spent there, and who would in the ordinary course of events soon be the university dean, but the time has long since passed when the idea that substantial parts of Himalayan palaeontology might be based on salted specimens, collected from one place and described as if they had been found in another. □

Minority of one

Britain seems determined to be perpetually at odds with all its fellow-members of the European Union.

Last weekend's meeting of the European Council, the governing body of the European Union (EU), was evidently disastrous. It was not the first of that kind, of course. If anything, disasters were more frequent when the EU was the European Community and when, in the early 1960s, almost every other European Council appeared to get nowhere. But last weekend's meeting on Corfu may presage the beginning of a long spell in which Britain dissents from the opinions of all its fellow-members on all issues that happen to arise.

It is a matter of political dynamics. The British government is sustained in office by a party deeply divided on Britain's membership in the EU. A substantial fraction of government Members of Parliament wish that the EU did not exist or, given that it does, that it would go away. That same group is large enough to ensure that the government would lose any vote in the House of Commons if it chose to vote with the opposition parties against the government. So the government is sustained in power by the restraint of its internal critics on European membership. What more natural than that they should use European issues repeatedly to test the government's compliance with their wishes? There is usually at least one potential quarrel on the boil. Long before last weekend, it had become plain that the government party's dissenters would make trouble if the Prime Minister of Belgium, Mr Jean-Luc Dehaene, were to succeed M. Jacques Delors as president of the European Commission. Britain's prime minister, Mr John Major, obliged his dissenters by saying that he would not have him.

That this development will cause trouble is easily foretold, but the form the trouble will take is not. Part of Britain's objection to Dehaene is that his name had been put forward by France and Germany without consultation with Britain. But Dehaene had also earned himself in Britain the self-contradictory reputation of being at once a "pragmatist" and a "federalist". The government could have tolerated the first, but its internal critics abhorred the second. It seems unlikely that Britain will be any more fully consulted on other candidates than in the first round.

That is a great misfortune, for the British position on the issues with which it has quarrelled with its EU partners in the past few years need not have been made to seem unreasonable. On last weekend's business, the British government would have performed a public service by pointing out that the power that has accrued during the past eight years (a measure of his competence) is anomalous. The European Commission is supposed to be the EU's executive branch, drafting legislation and looking after the money. The snag is that its members are appointed for fixed periods of time, and are awkwardly in between cabinet ministers and senior civil servants. The commission is an institution ripe for change. The British government would have done better to have said that at Corfu rather than saying it would not have Dehaene. □

Report prompts US to enter talks on new split of AIDS patent royalties

Washington and Paris. The United States this week opened talks with France aimed at ending a long-running battle over the joint US-French AIDS blood test patent. The move follows the completion of a two-year US inquiry which questions the claim of Robert Gallo of the National Institutes of Health (NIH) to have invented the US test.

Lawyers from the US administration and the Pasteur Institute were scheduled to start talks this week in preparation for a meeting of the foundation that administers the royalties on 11 July. Maxine Schwartz, director general of the Pasteur Institute, says he will seek an increased share of the royalties for his institute at the meeting.

For over two years, the US has been rejecting French pressure to re-negotiate an agreement that all royalties arising from the patent should be shared equally between the two countries.

But last week Harold Varmus, director of NIH, wrote to Schwartz agreeing that their lawyers should discuss the wording of "an acknowledgement from me appropriate to the current state of knowledge: that the French virus was used by NIH scientists in developing the American test kit".

Varmus did not concede that the Pasteur should get a bigger share of the royalties from the patent agreement. But he agreed with Schwartz that this would be straightforward. "Were I to be persuaded that a change in our current arrangement... is warranted, I would surely take steps to see that a change is made," the letter says.

A senior NIH official confirms that Varmus has shifted his position markedly since 8 June, when he wrote to Schwartz that "no alteration to our shared royalty arrangement is warranted", and asked him to "put the matter behind us". Varmus declined to comment on his correspondence with Schwartz, which NIH released in response to a Freedom of Information Act request.

The 1987 settlement confirmed that Gallo and Luc Montagnier of Pasteur had 'co-discovered' the virus. Under its terms the US Department of Health and Social Services (HHS) and Pasteur agreed to each donate four-fifths of the royalties from sales of tests to a French and American AIDS Foundation. A quarter of this pool then went to the World AIDS Foundation, while the remaining three-quarters was split equally between HHS and Pasteur.

Pasteur has fought for a bigger share on the basis that Gallo acknowledged in 1991 that the HTLV-III virus which he used to make the US test was in fact a virus isolated

previously by French researchers, including a group at Pasteur. Schwartz emphasizes that Pasteur explicitly stated that Gallo could not use the samples it sent to him for commercial purposes.

Varmus's change of position appears to have been prompted by conversations with Schwartz and by the completion earlier this month of a two-year inquiry into the Gallo affair by the inspector general of the HHS.

The inquiry, which covers the period 1983 to 1994, is broader in scope than one carried out previously by HHS's Office of Research Integrity (ORI). ORI dropped its efforts to convict Gallo of misconduct last year, after its case against his co-worker Mikulas Popovic collapsed on the grounds that it did not constitute legal proof of intent to falsify data, or that anyone had been defrauded (see *Nature* 366, 191; 1993).

Extracts of the summary of the new inquiry — which were first published in the *Chicago Tribune* — criticize Gallo's role in developing the AIDS test. It cites the patent examiner who granted the patent on Gallo's test in 1985 as saying she would not have done so had she known that the French also had a patent application in hand.

The summary claims that Gallo failed to disclose to the patent office that the French had performed "extensive work" on the virus, and that his laboratory had cultured it "for an extensive period and used it for



Gallo: claims no new allegations.

many of their experiments". NIH is taking the findings of the inquiry "extremely seriously" says the NIH official. Schwartz describes its conclusions as "damning".

Gallo says that it contains no new allegations against him, adding that, under an agreement with his employer, the National Cancer Institute, he cannot comment further in public. But his lawyer, Joseph Onek, has written to the inspector general rejecting the 10 June summary. He says it is "is filled with an extraordinary number of errors reflecting deliberate factual distortions, scientific illiteracy and obvious bias".

Renegotiating Pasteur's share of the royalties would not require rewriting the 1987 settlement, which does not mention how royalties should be distributed. A motion from the board of the French and American AIDS Foundation would be sufficient. Schwartz told Varmus this would be "a workable, simple solution that does no injustice to Dr Gallo, deprives him of no rights, but still rights a wrong".

HHS has received over US\$20 million in royalties since 1987, and Pasteur US\$14 million. Schwartz says he will seek changes in the distribution that would double the FF5 million (US\$850,000) royalties that Pasteur now receives annually. If he does not obtain a satisfactory solution, "then all options will be open" — including going to court.

But if Varmus makes concessions on the distribution of royalties this will inevitably re-open one of the most celebrated recent instances of alleged scientific misconduct in the United States, and raise new questions about the failure of either NIH or the wider scientific community to resolve it satisfactorily. **Colin Macilwain and Declan Butler**

Television appeal funds genetics institute

Munich. Money raised through a charity drive on Italian television is being used to build a new institute for human genetics in Milan. The institute, due to open next January, will be paid for entirely by Telethon, a charity supporting research into genetic disorders.

The new institute will be headed by Andrea Ballabio, currently at the Institute for Molecular Genetics, Baylor College of Medicine in Houston, Texas. Known as TIGEM (Telethon Institute of Genetics and Medicine), the new institute will be housed in laboratory space donated by the San Raffaele Biomedical Science Park, where the biomedical research centre DIBIT is already located

(see *Nature* 360, 7; 1992).

In line with Telethon's new policy of concentrating its research funding on large-scale projects, TIGEM will eventually have a staff of 35. Telethon will pay all administrative costs, as well as salaries, estimated at about IL1.8 billion (US\$11.5 million) a year. The laboratory will be set up with two grants of IL4.5 billion and IL3.5 billion for 1995 and 1996 respectively.

Research activities at the institute will include the positional cloning of genes, and identification of gene mutations leading to genetic diseases such as Kallmann's syndrome and ocular albinism.

Alison Abbott

US universities face spectre of cuts in military support

Washington. A key Congressional committee has called for US Department of Defense (DoD) funding for university research to be halved, from \$1.8 billion to \$900 million.

The House defence appropriations subcommittee, chaired by John Murtha (Democrat, Pennsylvania) voted for the draconian cut last week, citing the pain being suffered by the armed forces, and what Murtha sees as the threat of a "hollow force" with highly effective technology but no troops.

The proposed cut will doubtless be modified before the lengthy congressional budget process is completed in September. But the subcommittee proposal, which was approved in a closed meeting, set alarm bells ringing at prominent universities such as Stanford, Massachusetts Institute of Technology (MIT) and the University of Southern California, each of which receives about a quarter of its total federal research support from DoD.

Cornelius Pings, president of the Association of American Universities, which represents the biggest research universities, says that "the impact could be enormous, with some institutions losing tens of millions of dollars".

The universities have many powerful friends, especially in the Senate, and are optimistic that the cut can be reversed. But

congressional staff say that a more likely outcome is a compromise by Murtha, which will still leave the universities facing a substantial loss of income.

Congressional staff also say that one reason for the threatened cut is to fire a warning shot across the bow of House science committee chairman George Brown, who has been campaigning against the habit of appropriations subcommittees of "ear-marking" funds to pet projects (see *Nature* 366, 5; 1993).

In 1991, the last year for which figures are available, the largest recipients of DoD research funding were MIT, with \$55 million, Pennsylvania State University (\$30 million) and Stanford (\$29 million).

Johns Hopkins University received \$260 million from DoD to run the Applied Physics Laboratory near Baltimore, Maryland. But that is categorized as development work and, according to a university spokesman, is not thought to be subject to the proposed cut.

The Murtha subcommittee budget was passed on Monday by the full appropriations committee chaired by David Obey (Democrat, Wisconsin), and was due to be considered later this week by the House of Representatives as a whole. **Colin Macilwain**

Allain loses appeal and faces threat of new charges

Paris. A leading French haematologist who was given a four-year prison sentence (with two years suspended) for his involvement in a blood contamination scandal in the mid-1980s, last week lost his appeal to the Paris Supreme Court of Appeals.

The formulation of the court's ruling may also open the way to further charges of "poisoning" being brought against Jean-Pierre Allain, who formerly worked for the French National Centre for Blood Transfusion (CNTS).

Allain had appealed to the Supreme Court on the grounds that he should never have been convicted in 1992 under a 1901 law on "deception over the quality of the products he sold", after losing an appeal last summer in a lower court against that 1992 verdict.

The lawyer representing Allain, Olivier Schnerb, had also argued that, as Allain had worked in the research department of the CNTS, he had not been involved in distribution of contaminated blood products to haemophiliacs.

The court rejected Allain's appeals — the Supreme Court judges on the legal basis of a trial, and not the verdict — as well as 14 civil suits calling for Allain to be tried instead on the charge of poisoning, in a criminal court, along with three other officials convicted with Allain.

Of the three, Michel Garretta, the former director-general of CNTS, had received a four-year prison sentence and a FF500,000 fine on the same charge as Allain. The other two, Jacques Roux, former director-general of health at the National Ministry of Health, and Robert Netter, former head of the National Laboratory, had received three- and one-year suspended prison sentences respectively for "non-assistance to persons in danger".

Max Lecocq, a young haemophiliac, indicated after last week's appeal that he would bring poisoning charges against Garretta and Allain and "all the others". Lecocq has already brought charges of poisoning against three others who were ministers at the time, Laurent Fabius, prime minister, Georgina Dufoix, minister of social affairs, and Edmond Hervé, secretary of state for health (see *Nature* 367, 304; 1994).

Many in the judiciary, however, are likely to resist fiercely the idea that someone could be tried on different charges for the same act. But one thing seems certain — that Garretta and Allain will serve the full length of their sentences. Although most prisoners are released half-way through their term, the justice minister, Pierre Méhaignerie, said in a radio interview earlier this week that both should serve their full terms.

Declan Butler

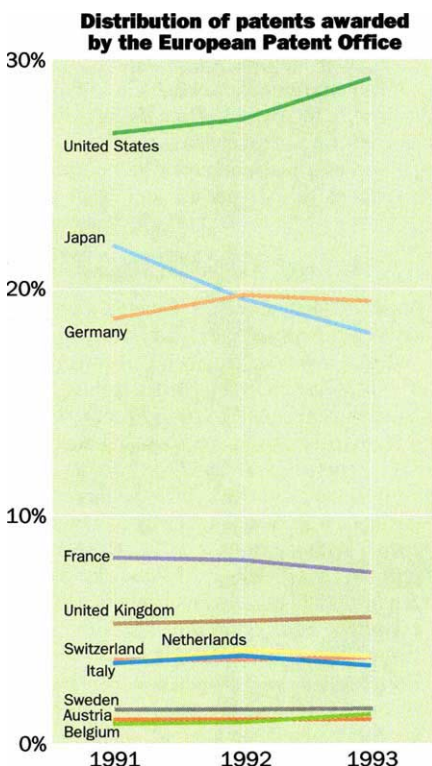
Recession hits European patents

Munich. To judge by patent applications received last year by the European Patent Office (EPO), the economic recession of the past few years has been taking its toll on innovation, with only the United States managing to register an increase in its innovative potential.

The overall number of applications to the EPO fell from its peak of over nearly 59,000 in 1992 to 57,000 last year. The United States filed 29 per cent of this number, compared to 27 per cent in 1992; in contrast, while Japan's share of the total dropped from nearly 19.5 per cent to 18 per cent.

The share of applications filed by the EPO's 17 member states fell from 49.3 per cent to 48.6 per cent, despite efforts by governments to beat the recession through national innovation policies.

Italy, France and the Netherlands fared particularly badly, each filing around 10 per cent fewer applications than in 1992, while even Germany's healthy record of 11,500 applications in 1992 fell by five per cent. The United Kingdom maintained its rather low share of around 3,100 applications (around 5.5 per cent of the total received by the EPO). **A. A.**



France, Switzerland 'must pay more' for LHC

Munich. Germany has thrown a spanner into negotiations over the construction of Europe's planned Large Hadron Collider (LHC) by unexpectedly demanding that France and Switzerland jointly provide an additional 10 per cent of the collider's costs, in addition to their regular contributions to the European Laboratory for Particle Physics (CERN).

Germany's demand dominated last Friday's meeting of the CERN council, which many had hoped would give formal approval to the construction of the LHC. In particular, negotiations on the contentious issue of inflation-linked rises in contributions seemed to be reaching compromise.

But the German delegation, acting under instructions from Bonn (and with the formal backing of the United Kingdom), demanded that CERN's two host countries, France and Switzerland, contribute an extra 10 per cent of the estimated SFR2.6 billion (US\$1.96 billion) costs of the LHC programme.

Neither country is likely to be in a position to raise a significant portion of such a figure, particularly in view of the pressures on France's science budget (see *Nature* 369, 511; 1994), and Switzerland's need to find extra funds in order to participate in the European Commission's Framework Programme (see *Nature* 369, 264; 1994).

CERN president Hubert Curien adjourned the meeting in order to give delegates more time to discuss the matter with their governments. Before doing so, however, he elicited a vote to proceed with the LHC from 17 of its member countries.

Curien hopes to reconvene the meeting before the summer holidays to allow Germany and the United Kingdom to vote. But he also accepts that negotiations may take until the next planned council meeting in September.

Meanwhile, he believes that the positive vote from the majority of countries will give CERN more weight in its negotiations with non-member states, such as the United States, Japan and Canada, which it hopes will contribute to the LHC programme. This currently has a shortfall of SFR500 million in its budgeted funds.

The issue of whether France and Switzerland should pay a significant premium for the privilege of being host countries to CERN (the laboratory straddles the border between the two countries) has been raised on previous occasions. The idea has wide support among member states, who realize the value of the laboratory to its two local economies. As countries such as Spain have pointed out, France enjoys a financial return, measured by the total of value of contracts won by French enterprises, more than double its contribution; Switzerland's return is six times as much.

Switzerland has in general been sensitive to this issue, and has made several *ad*

hoc payments over the years to contribute to specific programmes. It announced in December that it was prepared to make a significant contribution above its normal subscription to the building of the LHC.

But, to the annoyance of Germany — which feels it has shouldered a large portion of CERN's costs in the past — France has held back. Only in the past few weeks has it made a comparable statement.

Neither the Swiss nor the French offers have figures attached, though current estimates suggest neither is much above SFR25 million. Both countries appear to be waiting as long as possible to see how far non-member states are prepared to meet the expected budgetary shortfall.

The sudden demand from Germany that the two countries should pay between them a SFR260 million premium over ten years received mixed responses from other member states, frustrated that Germany appeared to be spoiling an opportunity to reach consensus on LHC.

There were also murmurings about Germany's actions being dictated by a higher political agenda, such as rumoured attempts to link the CERN issue to other political decisions. But most accepted that it is primarily Germany's tight budget situation, a result of reunification, which prevents it from being as generous as previously.

There was also some feeling of support for the way that Germany criticized France for not paying a premium on its contributions to CERN in recognition of the country's host status.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Llewellyn Smith:
seeking compromise.

Hermann Strub, the German delegate, refused to comment on Germany's tactics after the meeting. But he claims that a 10 per cent premium for host countries of international laboratories has become the norm, citing the Joint European Torus (JET) in Britain, and the European Synchrotron Radiation Facility (ESRF) in France.

Meanwhile a second issue, which had threatened to delay approval of the LHC, seems closer to resolution. Germany and the United Kingdom have both been keen to hold down annual inflation-linked rises in contributions by changing voting rules so that such rises could be vetoed by any of CERN's 19 member states. (see *Nature* 369, 509; 1994). At present, a rise can be agreed by a simple majority of member states, provided their contributions represent at least 55 per cent of the budget.

CERN management refused to accept that this should now require unanimous approval. It pointed out that it would allow any country, no matter how small its financial contribution, to block a rise in budget, seriously reducing CERN's buying power.

Negotiations in the last few weeks between CERN director-general Christopher Llewellyn Smith and Germany and the UK had succeeded in moving towards a compromise whereby contributions would be fixed at 1994 levels until 1997, and then reviewed at the time of the major LHC programme review.

In addition, a compromise voting procedure, suggested by Belgium, which would require a majority representing around two-thirds of the budget, is under consideration. The figure finally arrived at is critical, as Germany and the United Kingdom together contribute around 36 per cent of CERN's budget, and could combine to block a rise. That makes some countries, such as Switzerland, uneasy, but a compromise between these positions seems within reach. **Allison Abbott**

Britain argues for 'firm cost control'

London. Ken Pounds, professor of astronomy at the University of Leicester and chief executive of Britain's newly-formed Particle Physics and Astronomy Research Council, said on Monday that it was essential for Britain to put strong financial constraints on the costs of the planned Large Hadron Collider (LHC).

Speaking at the opening of an international conference on particle accelerator technology, Pounds said that in recent years funding available to British scientists to work on the results of experiments at CERN had fallen by a factor of three. "This is like paying high fees to join a golf club, and then having no money left to buy any balls."

Pounds, who expressed optimism that

the French would agree to make a "substantial" additional contribution along the lines being demanded by Germany (see above), said that Britain is keen for the LHC to proceed. "But we felt this was the moment to say: hang on."

David Davies, parliamentary secretary in the Office of Public Service and Science, said that the UK was convinced that the LHC was "the right way forward" for high-energy physics. But, reflecting Britain's desire to see CERN enlarged into a "fully global organization", he said that large-scale facilities such as the LHC "cannot — and indeed should not — be afforded on a national or even a regional basis".

M. V.

Booster test lifts prospects for new launcher

Kourou, French Guiana. The European Space Agency (ESA) is on course to launch its first Ariane 5 rocket in October next year, following the successful third ground test of the rocket's solid fuel booster, carried out last week at the agency's space centre in French Guiana.

The P-230 booster — one of the pair that, when strapped to Ariane 5's central Vulcain cryogenic motor, will provide 92 per cent of thrust at lift-off — is ten times bigger than any rocket previously built by Europe. At 31 metres high and 3 metres in diameter, each booster alone provides more power than the Ariane 4 launcher now in service.

In 133 seconds, the booster burned 237 tonnes of butadiene, aluminium and ammonium perchlorate propellant. About 230 tonnes of exhaust gases formed a huge white cloud that mushroomed to 11,000 feet above the forest. The firing also pulverized up to 150 tonnes of granite torn from the top 50 centimetres of the rock below the test-bed.

The French space agency CNES, which has responsibility for the tests, admits that the control of pollution from the exhaust is a problem, and restricts tests to 'appropriate' meteorological slots. To prevent the 48.3 tonnes of hydrochloric acid in the cloud from falling on trees, for example, CNES requires 2.5 hours without rain after a test.

Held tightly by an ECU50-million gantry, the booster lifted just 5 centimetres off the ground, sufficient for certain tests to be carried out. First analyses of roughly 600 different measurements suggest that the booster performed as predicted. In particu-

lar, it produced the desired flight thrust profile, including a surge to 540 tonnes in the first 15 seconds.

Development of the solid-fuel boosters accounts for almost a fifth of the ECU6.1 billion cost of the Ariane 5 programme. New industrial ground facilities in Guiana for manufacturing the fuel and assembling the rocket have cost a further ECU910 million; ESA decided that production *in situ* would be safer and simpler than transporting the boosters from Europe.

CNES will test the booster four more times this year, and Ariane 5 will have its first test flight in October 1995, with a subsequent test in April 1986. ESA will test a small unmanned reentry vehicle attached to the launcher's nose-cone on one of these flights. The first commercial flight of Ariane 5 is scheduled for October 1996.

The Ariane 5 programme was approved by ESA members' space ministers at La Hague in 1987, and is planned to give Europe a more powerful and more reliable launcher that will be 10 per cent cheaper than Ariane 4. France has agreed to pay 46.2 per cent of the costs, with Germany covering 22 per cent and Italy 15 per cent.

On the basis of market research, the agency predicts that satellites weighing be-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Up to 150 tonnes of granite were pulverized in the test.

tween 2.4 and 3.6 tonnes will soon make up three-quarters of the commercial launch market. Arianespace, which now takes half of this market, needs the new launcher to compete with the Russian Proton and Chinese Long March-3 launchers.

Ariane 5 is designed to place two satellites with a mass of 5,900 kg — or one of 6,800 kg — into a geostationary transfer orbit (36,000 km), 10 tonnes into Sun synchronous orbit (900 km), or 22 tonnes into low Earth orbit (300 km). ESA is also planning bigger and smaller versions of Ariane 5.

Designed also to carry the now-abandoned Hermès space shuttle, Ariane 5 should also be sufficiently reliable to give Europe an independent manned flight capacity.

In particular, ESA recently approved a ECU207-million manned programme, including an Apollo-like capsule to be mounted on the nose of Ariane 5. It may also use Ariane 5 to service the international space station, if the US-led project goes ahead.

To achieve its goal of cutting the costs of Ariane 4 by 10 per cent, Ariane 5 will need to make eight flights a year, including five commercial ones. Many think Ariane 5 will make this target, but perhaps not before early next century.

Costs have overrun by 25 per cent, and the Paris-based Euroconsult space consultants say the satellite market may dip between 1996 and 2000. The expected regular Hermès missions will also not now materialize, although space station missions would compensate for this loss.

For its part, Ariane 4 flew two weeks ago for the first time since January, when a failed launch lost satellites worth US\$356 million (see *Nature* 367, 398; 1994). Arianespace, which has contracts to launch 40 satellites on its order book worth FF18.7 billion, is speeding up its launch schedule to make up for this five-month delay, required to modify the liquid oxygen/hydrogen turbopump which caused the accident. **Declan Butler**

FDA claims nicotine levels were manipulated

Washington. David Kessler, commissioner of the US Food and Drug Administration (FDA), last week presented to a congressional subcommittee "conclusive evidence" that the tobacco industry deliberately manipulates nicotine levels in cigarettes.

Kessler told the health and environment subcommittee of the House Committee on Energy and Commerce that the cigarette manufacturer Brown and Williamson had deliberately developed seeds to produce high-nicotine, low-tar tobacco plants.

Kessler also told the subcommittee, chaired by Henry Waxman (Democrat, California), that at least one other company puts ammonia in cigarettes. The ammonia increases the nicotine levels received by smokers by binding with the salts in tobacco.

"These findings lay to rest any notion that there is no manipulation and control of nicotine undertaken in the tobacco industry," asserted Kessler. The question of whether or not the industry deliberately manipulates nicotine levels is crucial to

whether the FDA has the legal right to regulate nicotine as a drug.

The FDA's investigation of the tobacco industry and the slew of congressional hearings during the past few months have been aimed at building support in both Congress and the public for the regulation of cigarettes. Advocates of regulation must win enough votes in Congress to defeat the powerful tobacco lobby.

Two days after Kessler's testimony, Thomas Sandefur, the chief executive of Brown and Williamson, denied the FDA allegations and told the subcommittee he did not believe nicotine to be addictive in the sense that cocaine and heroin are addictive.

The question of addiction will be discussed in August at a meeting of the FDA's drug abuse advisory committee. The relative addictiveness of nicotine compared with other drugs such as cocaine will inevitably arise, as many members of the panel are experts in the treatment of addiction to these illegal drugs.

Helen Gavaghan

Chinese scientists to get Western form of recognition

London. Chinese science has moved further towards remoulding itself along Western lines with a decision to designate the members of the scientific council of the Chinese Academy of Sciences as "academicians".

The academy has also announced its first 10 foreign academicians, including the eminent British historian of Chinese science, Joseph Needham. And a companion body has been set up, the Chinese Academy of Engineering, as a focal point for the development of new technologies.

The science academy's scientific council is China's leading science and technology advisory body.

In the past, China has formally eschewed the élitism implicit in Western scientific institutions by refusing to give any extra recognition to members of the council.

According to academy officials, the reason for granting council members the title of academician is to facilitate international exchanges and express the "authoritative and honorary status" of the council. A similar system will be introduced for the engineering academy.

Speaking at a meeting of both bodies in Beijing earlier this month, the Chinese premier, Li Peng, described the move as "an important measure which conforms to the practice of the scientific and technological communities of the world, as well as the strong and general aspirations of the scientific and technological communities at home".

He also said that the founding of the engineering academy is a major event for China's engineering communities and an important part of the country's drive for modernization. The appointment of the foreign academicians would have a "positive impact on opening China's scientific and technological field wider to the outside world", he added.

David Foster of Britain's Royal Society said that the society welcomed the decision to describe the scientific council members as academicians. It gave "a clear indication of the desire of the Chinese government to continue to accord scientists and engineering a high status which matches that of their counterparts in other countries".

He also described the decision to elect the first foreign academicians as "a landmark in the history of scientific links between China and the rest of the world". □

Deaths bring clinical trials under scrutiny in Japan

Tokyo. Deficiencies in Japan's procedures for monitoring and assessing clinical trials have been brought to light by a rapidly developing scandal over a new antiviral drug that has killed several patients, both during trials and after receiving government approval.

In the latest twist in the scandal, the Securities and Exchange Surveillance Commission (SESC) last week raided the head office of Nippon Shoji Kaisha, the company that developed the drug Sorivudine, on suspicion that employees have been engaged in insider trading.

The SESC's action was prompted by the fact that the employees sold shares they held in the company shortly before a government announcement on 12 October 1993 that several patients had died after using the drug in combination with anticancer agents.

After several years of clinical trials starting in 1985, Sorivudine, which contains an arabinofuranosyl uracil derivative as its active ingredient, was approved by the government last July for treatment of shingles, the adult form of chicken pox caused by the virus herpes zoster.

The drug went on the market on 3 September. But within days, several deaths were reported in cancer patients being treated for shingles when the drug was used in combination with fluorouracil anticancer drugs.

The company suspended shipments of the drug on 12 October, after the government had announced the first deaths. But by November, 15 deaths, and numerous cases of serious side-effects, had been reported.

Two weeks ago, *Asahi Shimbun*, one of Japan's leading business newspapers, revealed that three people had died during clinical trials when the drug was used in combination with anticancer agents. But the company reported only one death to the Ministry of Health and Welfare when it submitted the drug for licensing approval.

Sorivudine is thought to act with fluorouracil to inhibit the anticancer agent's breakdown in the body, making it more toxic. Sorivudine carried a warning that it should not be used in combination with anticancer agents. But the warning did not specify that death could result.

Masanori Fukushima of Aichi Cancer Centre, who has been campaigning for the revision of Japan's clinical trial system (see *Nature* **342**, 850; 1989) describes the Sorivudine case as "shameful", but says it is "by no means an exception". He and other critics point out several defects in the clinical trials of the drug.

For example, reports on the trials were published in Japanese only in local journals, and the quality of many of the papers in

these publications is "very poor", says Fukushima. The abstract of the paper on the phase III clinical trial makes no mention of the number of patients, or of the side-effects (including death) that occurred.

More serious are the accusations that some of those who carried out the clinical trials should have been aware of the drug's potentially dangerous interaction with anticancer agents. Animal studies in Belgium, Japan and the Netherlands published in 1986 and 1987 indicated that the drug bromovinyldeoxyuracil (BVDU), which has a similar structure to Sorivudine, increases the toxicity of fluorouracil anticancer drugs.

M. Niimura of Jikei University School of Medicine, who supervised the Sorivudine clinical trials, is reported in *Asahi Shimbun* as saying that he read the Belgian BVDU study, but failed to link it to deaths in the early stages of the Sorivudine clinical trials.

Fukushima claims that a basic problem is the lack of good education in clinical pharmacology in Japanese medical schools.

A thorough investigation is still awaited of the reasons why the government's approval process failed to pick up the dangers of Sorivudine. One reason, says Fukushima, is inadequate expertise in the ministry's drug approval committee.

Another reason, he suggests, is the lack of an adequate independent system for inspecting clinical trials. In the United States the Food and Drug Administration has more than a hundred inspectors to check the conduct of such trials; Japan has only two.

David Swinbanks

Malaysia to boost pay for foreign scientists

London. The Malaysian government is drawing up proposals to offer Western salaries to foreign scientists and engineers to attract them to work in the country. The offer, which is expected to form a key element of the government's plans to become a fully industrialized nation by the year 2020, will also be open to Malaysian scientists and engineers currently working abroad.

According to the Bernama News Agency in Kuala Lumpur, a paper drawn up by a working party from the ministry of science, technology and the environment setting out the details of this plan are currently awaiting the approval of various other government departments, including the Treasury. If the endorsement of these agencies is received, the plan will be submitted to the cabinet for final approval. □

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Needham: an early
foreign 'academician'.

Needham Research Institute

Groundwater flow fans fears over UK nuclear waste store

London. New doubts have arisen over the safety of locating the United Kingdom's deep underground repository for radioactive waste at the proposed site at Sellafield, in Cumbria. The Radioactive Waste Management Advisory Committee (RWMAC), which advises the government, fears that contaminated water from the proposed repository site could rise to the surface, polluting surrounding land and water.

In a report published last week, RWMAC claims that its review of hydrogeological data from the area shows there is "inconsistency" between observed groundwater behaviour and results from computer-based modelling, which predict that water from the site would flow out into the Irish Sea.

The computer modelling has been carried out by UK Nirex Ltd, the nuclear industry's waste disposal company charged with selecting and constructing a repository site. The company has so far drilled more than 10 deep boreholes in West Cumbria for data collection using techniques that RWMAC acknowledges to be at the "leading edge" of international experience.

But in its annual report, RWMAC highlights what it sees as inadequacies in Nirex's conceptual modelling, focusing on the waste company's prediction of groundwater flow around the site.

According to RWMAC, data from the latest Nirex report (Report No. 525) indicate an upward movement of groundwater from the proposed repository site, which is located within the Borrowdale Volcanic Group, as it moves west through the Sherwood Sandstone Group (SSG) towards the sea. Groundwater is discharged to the surface from the SSG near the coastal margin (see diagram).

But the pathline predicted by one of Nirex's fracture flow computer models, NAMMU (shown in black on the diagram) predicts that water from the repository site will make a sharp decline along a fault line,

against the upward flow of the groundwater. Given that this would mean lower density fresh water moving downwards into a higher density brine, the interpolated pathline becomes "difficult to conceive", concludes RWMAC.

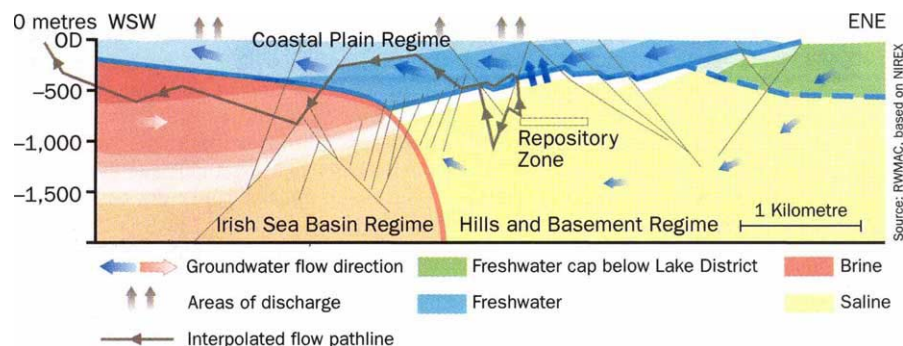
"As a result, the discharge of groundwater from the repository zone to the land surface above the repository site is no less plausible than the flow models illustrated by Nirex," says the report. It calls on Nirex to reconcile the differences between observations and computer-modelling predictions through model validation and data verification.

"We will ultimately see a repository built in Britain. Whether it is at Sellafield is a different matter. Clearly there are major questions that need answers," said Sir John Knill, chairman of RWMAC, at the launch of the report.

Tom Curtin, a spokesman for Nirex, points out that the company has still to collect all the information it needs, and expects to take at least three or four more years to do so. "If we had all the information right now we'd apply for planning permission for the repository itself," said Curtin. Instead, Nirex intends to seek planning permission in the next few weeks for the construction of a rock characterization facility (RCF) near the repository site for further investigation into groundwater flow. The RCF, costing around £120 million in addition to the £300 million that Nirex has already spent studying the site, is seen as "vital" to the process.

If all goes according to plan, work on the RCF will begin in 1995. Settling the question of a long-term waste site represents "a key policy question", according to another report published last week by the Parliamentary Office of Science and Technology (POST) on the government's long-awaited review of the prospects for the nuclear industry, announced last month.

Maggie Verrall



RWMAC's analysis suggests that Nirex's predictions of how water will move may not be born out in practice.

Fossil inquiry finds Indian geologist guilty of plagiarism

New Delhi. An inquiry by Panjab University in Chandigarh has apparently found the prominent Indian geologist Viswa Jit Gupta guilty of plagiarism, of recycling fossils and of claiming to have made discoveries in localities he had not visited.

The inquiry, prompted by concerns over Gupta's research that were first raised by the Australian geologist John Talent (see *Nature* 338, 613; 1989), also claims that Gupta was guilty of misleading his co-authors. The university is now deliberating on what action to take against Gupta, who still retains his position as professor at the university.

The report of the inquiry officer, M. S. Gujral, has been submitted to the university's vice-chancellor, T. N. Kapoor, and has been circulated among members of the 'syndicate', the university's governing body.

In particular, Gupta had been charged with claiming that fossils collected from museums and other sources were his own samples recovered on field trips, and that the many scientific publications based on these fossils had distorted the literature on Himalayan geology.

Two earlier investigations — one by the Geological Survey of India and another by the Society for Scientific Values (SSV) — had independently found Gupta guilty. But it was only after the cancellation of funds from the University Grants Commission and adverse publicity in the international scientific press (see *Nature* 355,660; 1992) that Kapoor decided to act.

The inquiry covered not only Gupta but also his four co-authors, all of whom had sided with Talent. Although the report is said to confirm Talent's charges against Gupta, it exonerates his co-authors of being party to the fraud.

Despite the delay — the inquiry started in February 1992 — the final verdict has been generally welcomed by the scientific community. "I am very, very happy it is over," said A. S. Paintal, who carried out his own investigation of the charges against Gupta as president of SSV in 1991.

Gupta defended himself against the charges, but apparently failed to produce any evidence sufficient to refute the charges. He was unable to remember the dates of his expeditions, bluntly told the inquiry that he did not keep field notes, and retracted a document submitted earlier to Gujral detailing a field trip to Khemkul La.

India's geologists are keenly awaiting the syndicate's decision on Gupta's punishment. Some are demanding firm action. "We cannot let Gupta off mildly," says Paintal. "He must go." Gupta himself has not been available to comment on the inquiry's conclusions.

K. S. Jayaraman

Rome 'could host European mouse laboratory'

Munich. The council of the European Molecular Biology Laboratory (EMBL), based in Heidelberg, Germany, was to be briefed this week on a proposal to construct a mouse genetics laboratory in Rome that could, its supporters hope, persuade Italy not to leave EMBL at the end of the year.

Italy announced its intention to quit last year, saying that it is dissatisfied with the returns on its membership subscription. Italy pays 16 per cent of the laboratory's annual budget — currently DM10 million (US\$6.3 million) — but there are very few Italian staff scientists in Heidelberg (see *Nature* 366, 604; 1993).

The project under consideration is an initiative backed by the European Commission (EC) in Brussels to set up a laboratory to act as a European-level repository for mouse mutants. As such, it would be complementary to the Jackson Laboratory at Bar Harbor, Maine, in the United States.

The idea for a similar European facility was proposed last year by Peter Gruss, a German cell biologist who is one of 25 personal scientific advisers to Antonio Ruberti, the commissioner for research. Gruss says that Europe urgently needs a facility to maintain stocks of the ever-increasing number of mutant strains, now being created at a rate of around 3,000 a year.

His idea has been enthusiastically supported by Ruberti and his advisory board, who have described it as "vital for studies of the diseases (such as cancer and coronary heart disease) that are the main causes of illness and death in the [EU]" (see *Nature* 369, 11; 1994).

Gruss argues that such a service laboratory would need to be combined with high-level research in order to maintain scientific standards. This is where EMBL comes in. Gruss proposes that EMBL should run the facility, combined with an active research team, as an outstation of its laboratory in Heidelberg.

At present, the idea is only a recommendation as far as the EU and EMBL are concerned. Before it can proceed, the EU's next biotechnology programme (within Framework IV), which includes funding for genetic repositories in Europe, must be approved. Only then can a call for proposals be made.

Italy is already working on its own proposal. Umberto Colombo, the former research minister, has written to his successor, Stefano Podestà, suggesting that urgent consideration be given to offering research laboratories at Monterotondo near Rome to EMBL as part of a bid to host the facility.

The laboratories which will provide the 20,000 square metres required for the proposed facility were built by ENICHEM, the loss-making chemical subsidiary of the state-owned Eni energy and chemicals group, but never put into service because of

ENICHEM's financial problems.

Italian scientists have asked Podestà for a meeting to discuss the proposal, and are awaiting a reply. Meanwhile EMBL must decide whether the proposed facility would fit in with its own long-term plans. The laboratory has declined to comment on the proposal, but has already sent a delegation to examine the site.

But EMBL's member states are likely to be positive, particularly because the proposal could bring Italy back to EMBL membership. Colombo predicts that Italy would

definitely rejoin if the mouse facility, backed by EC funding, is agreed, although the final decision on membership would be made by Podestà.

But Gruss, who is also one of Germany's delegates to EMBL, says that he would insist to the council that Italy agrees to rejoin before the idea is discussed further. The clock is ticking — EMBL must settle its 1995 budget by the end of the year and would very much like Italy to commit itself to coming back into the fold before then.

Allison Abbott

Future brightens up for California

San Francisco. With growing signs that retail spending and the construction industry are pulling California out of its three-year recession, the state's science-based industries offer a solid platform for long-term growth, according to a report just published by a group of economists.

But the economists also say that such growth will only take place if the state continues to invest in education and infrastructure, and could be jeopardized if insufficient funds are provided for state-run universities. Over the past three years, the state has cut the University of California's budget by 16 per cent.

According to the Center for Continuing Study of the California Economy in Palo Alto, foreign trade, high technology, professional services, and tourism and entertainment promise to provide the state with a healthy economic base and above-average growth prospects.

Even during the recession, the report points out, California continued to produce a record number of high-technology start-up companies. In 1993, for example, venture capitalists invested \$950 million in new Silicon Valley companies. Initial public offerings also hit a record, and profits at the 100 largest companies in Silicon Valley grew by 36 per cent.

Steven Levy, director of the centre and the report's principal author, says four factors prolonged the recession in California: a sharp decline in residential building, cuts in defence spending, weakness in commercial aircraft orders and a large relative decline in consumer spending.

But the construction industry predicts a good year ahead, and retailers have reported four consecutive strong months. Although 161,000 jobs were lost in the aircraft, space and defence-related industries in the past four years, says Levy, many of the aerospace job losses resulted from civilian market adjustments, and those jobs will return as aircraft exports rebound under strong demand.

The report predicts a reduction in the rate at which new jobs are created in the 1990s. But it says that the state will still continue to expand faster than the nation as a whole. Almost all of the California job growth would take place in non-manufacturing activities, many of which offer high wages, the report said.

But these opportunities will not be fulfilled unless the state develops a strat-



egy to fund critical public investments, warned Robert Arnold, co-founder of the economic think-tank.

The state now invests 14 per cent less than the national average in educating each public school pupil, according to the Center of Budget and Policy Priorities in Washington. California's spending on roads, bridges and other public works is also below the national average.

Arnold says that the state must maintain a skilled labour force, world-class physical infrastructure, a high quality of life and a competitive regulatory environment to attract private investment in key sectors. An excellent university system and highly regarded laboratories are critical to attracting and maintaining growth industries, he said.

Sally Lehrman

Prevention rather than cure

SIR — An innovative precautionary principle is emerging in international environmental policies of northwest Europe and Scandinavia¹. The precautionary approach is encouraging new thinking through its emphasis on clean production methods and prevention of environmental contamination in the belief that it is cheaper, easier and more practical to prevent pollution in the first place than to try to clean contaminated systems later on². As M. MacGarvin observes³, the precautionary ideal arises from recognition that scientific understanding of ecosystems is complicated by a host of factors, including complex and cascading effects of human activities and uncertainty introduced by naturally chaotic population dynamics.

Precaution also serves as a progressive policy tool. The principle addresses a key question for environmental managers: how should policies be decided in the face of scientific uncertainty? The response from science is to engage in further rigorous studies better to understand the hidden workings of nature. But a similar response is not available within the culture of policy; in a setting that must cope with demands for economic growth, the pressures for resource extraction are immense. So important policy decisions (including continuing the status quo) are made despite poor knowledge of the ultimate effects of anthropogenic activities; this situation is proving increasingly problematic for modern environmental management.

Precautionary thinking counters traditional regulatory emphasis on 'end of pipe' pollution control technologies. Moreover, precaution fundamentally rejects assimilative capacity — for instance of the coastal oceans — as a convenient means of (hoped-for) dilution. Vague definitions of precautionary action are evolving as it is increasingly applied. Initially the principle was put forward in an international setting at the first ministerial conference on North Sea pollution in Bremen in 1984 (ref. 4); it was strengthened at the second North Sea conference in 1987 (London) and further reinterpreted at the third conference in 1990 (The Hague)¹⁻⁴. The approach now appears in non-marine applications such as regulating pesticides in Africa (Bamako Convention, 1990), and efforts to reduce atmospheric contamination (Second World Climate Conference, 1990)⁵.

In sum, the emerging precautionary principle could become an important new link between science and policy. When first suggested, precautionary action was intended mainly to reduce marine contamination due to synthetic chemicals and heavy metals, and to halt dangerous activities such as ocean incineration because

the burden of suspicion made it prudent to prevent such activities³. In only a few years, preventive thinking behind this principle has spread to broader matters², and may yet become an eloquent step in achieving greener paths of development. Significantly, the United States has so far resisted the precautionary concept¹; whether the US position evolves towards acceptance of precautionary action will surely have much to do with the eventual fate of this rising principle.

Robert J. Wilder

*Department of Political Science,
University of Massachusetts, Dartmouth,
N. Dartmouth, Massachusetts 02747, USA*

1. Stairs, K. & Taylor, P. in *International Politics of the Environment* (eds Hurrell, A. & Kingsbury, B.) 110–141 (Oxford University Press, 1992).
2. Dethlefsen, V., Jackson, T. & Taylor, P., in *Clean Production Strategies: Developing Preventive Environmental Management in the Industrial Economy* (ed Jackson, T.) 41–62 (Lewis, Boca Raton, 1993).
3. MacGarvin, M. *Helgolander Meeresunter* **49**, in the press (1994).
4. Gundling, L. *Int'l J. estuarine coastal Law* **5**, 23–30 (1990).
5. Weintraub, B. A. *N. Y. U. envtl. Law J.* **1**, 173–223 (1992).

Malaria and AIDS

SIR — K. S. Jayaraman (*Nature* **378**, 90; 1994) contrasts planned spending in India of US\$27 million for AIDS control with the \$2 million for the control of malaria.

But while malaria is treatable, AIDS is not. If AIDS cases increase as some have predicted, India could end up spending Rs 34,500 crore (\$11 billion) by 2000 on one million AIDS patients. India's commitment to the control of malaria is demonstrated by the allocation of \$68 million for 1994–95 by the central and state governments, as well as additional spending by local governments, hospitals, industries and research organizations. The budget of the Malaria Research Centre for 1994–95, for example, is \$1.2 million.

Money is not lacking, but technical and operational problems need to be tackled. In rural areas, residual spraying with insecticides is ineffective. Other problems include lack of mobility of health workers, poor supervision, delay in supplying insecticides, inability to pay wages, demoralized staff and frequent transfers.

The result is a lack of understanding of the different types of malaria, some of which are appearing in areas previously free of the disease. In some places, malaria does not respond to DDT, and reservoirs of the disease remain. Other areas have drug-resistant strains.

In short, malaria is getting out of control. What is needed is local knowledge of the dynamics of malaria transmission together with sound knowledge of the biology of vectors to produce an appropri-

ate mix of technologies with which to combat the disease. Unfortunately, however, knowledge about malaria is diminishing in countries where it is endemic.

Knowledge, not money, is the key to defeating malaria.

V. P. Sharma

*Malaria Research Centre,
22 Sham Nath Marg,
Delhi – 110 054, India*

Vatican error?

SIR — If all members of the new Pontifical Academy for Life "must be prepared to acknowledge that life begins at the moment of conception" (*Nature* **368**, 90; 1994), one must agree with Gerard F. LeBlond (*Nature* **369**, 352; 1994) that this is indeed problematic. That is because, although both ovum and spermatozoa are certainly alive, destroying them before the egg is fertilized must be morally neutral.

But what the Vatican presumably meant was that it is a new genetically unique living organism that begins its life at the "moment of conception" and that, because this can be expected to complete its full development unless it dies, then intentionally to destroy it does present ethical and theological problems. That is all the more the case because for several days after conception one cannot say that a new *individual* life has begun, because there is still a possibility that the fertilized egg may develop into genetically identical twins.

C. B. Goodhart

*Gonville & Caius College,
University of Cambridge,
Cambridge CB2 1TA, UK*

Blinkered review

SIR — In his somewhat blinkered review of Kitty Ferguson's *The Fire in the Equations* (*Nature* **369**, 198; 1994), Frank Tipler makes the absurd claim that scientists could "only take God-talk seriously" if it led to statements such as "If God exists, the top quark must have a mass of 185 ± 20 GeV". It would be pretty odd to decline to take a physical theory seriously if it did not lead to so specific a prediction. To apply the criterion to the evaluation of a metaphysical understanding involving questions of meaning and purpose (for that is what theism is about) is just an elementary category mistake. Tipler's remark is as illogical as declining to take quantum chromodynamics seriously because it tells us nothing about the significance of human existence.

John Polkinghorne

*The President's Lodge,
Queens' College,
Cambridge CB3 9ET, UK*

Wrong committees axed?

SIR — A recent leading article (*Nature* 368, 482; 1994) praised the termination by the US Office of Management and Budget of a large number of scientific and technical advisory committees. On the face of it, such a move is laudable, but there are deeper implications.

For one thing, the approach taken to reducing the number of advisory committees was arbitrary and heavy-handed: the National Institutes of Health (NIH) and the Food and Drug Administration (FDA) were told to reduce their advisory committees by one-third, regardless of their merits. Many of them make a vital contribution: at FDA, they speed up and lend certainty to the regulatory review of pharmaceuticals, and at NIH they provide essential peer-review of the arcana of grant proposals and other programmes (*Science* 264, 191; 1994).

By contrast, panels that are clearly superfluous but 'politically correct' have been spared; NIH's Recombinant DNA Advisory Committee (RAC), which has evolved into nothing more than a *de facto* regulatory agency for human gene therapy proposals, is one example. On its merits, the RAC should go. First, the NIH's jurisdiction provides three additional levels of review (local institutional biosafety committees, the RAC and the NIH director) that are internally redundant and also duplicate the legally required evaluations of institutional review boards and the FDA.

Second, the RAC lacks expertise relevant to certain aspects of the regulation of pharmaceuticals, including manufacture and quality-control testing, with the nuances of clinical trials. Membership is heavy on lawyers, political scientists and ethicists.

Not surprisingly, the NIH has sometimes held the protocols to standards very different from those the FDA applies to all other experimental treatments. The result has been that researchers are confused and spend their time on paperwork instead of experiments.

Third, the NIH committee holds only a few brief meetings a year, forcing investigators to wait several months for a proposal to be considered. Moreover, because the RAC is not continuously available to them, researchers are unable to obtain permission to make important real-time adjustments to the experiment, if problems occur or if other developments dictate changes in course. The NIH committee has recently made some modest recommendations intended to streamline its regulatory review, but even if its oversight were substantially more efficient, the committee's worth would still founder on the principle that something that is not worth doing at all is not worth doing well.

Especially when the administration is applying the "reinventing government" axe to important advisory panels at the NIH and at the FDA, eliminating unnecessary layers of gene therapy regulation would be a win-win proposition: reduction of unnecessary federal and researchers' spending on regulation; more resources available for the actual research; and with diminished regulatory disincentives, greater interest in gene therapy from industry. Improvements in the trim and motion of the ship of state are important, but politicians must be circumspect about which and how much ballast to discard.

Henry I. Miller

Hoover Institution & Institute
for International Studies,
Stanford University,
Stanford, California 94305-6010, USA

ORI should stay

SIR — In the best of times, it would be gratifying that there remain adherents to the Platonic ideals; the existence of an absolute good comes to mind. In that spirit, I wish to dissent from the opinions expressed in your leading article regarding the necessity (or lack thereof) for an Office of Scientific (now Research) Integrity (ORI) (*Nature* 368, 1–2; 1994).

While it is fabrication and falsification of research results that receive most attention, particularly in the popular press, plagiarism is the more egregious of common sins, particularly in research. Plagiarism is often punishable by expulsion from school or university, but, if it goes undetected, may become institutionalized and accepted (within certain confines). If plagiarism takes place in an institution of higher learning, particularly one with a relatively decentralized bureaucracy, there may be a (retaliation-free) recourse for the subject of the plagiarism. But if such an offence occurs in more 'para-academic' institutions (museums and research institutes) wherein the hierarchical systems may be more rigid (oligarchical at times) and without the habitual academic freedoms some take for granted, the individual subject to the offence may have no such recourse.

I would therefore submit the following: your first lesson, that the ORI is not suited to the task for which it was created, is without doubt true. Your second lesson, that the ORI should be abolished, does not follow. Indeed, its mission should be expanded. At present, only institutions receiving federal funds are required to establish policies and procedures for handling misconduct allegations. Furthermore, it is the institution in which the

misconduct occurred that is responsible for undertaking the investigation. And even when misconduct occurs in an institution receiving federal funds, if the particular instance of misconduct does not directly involve a definite federally funded project, then there is no recourse to the ORI. In these times of diminishing (or at least harder to obtain) federal research monies, an increasing number of potential misconduct cases are consequently free from any oversight besides the perhaps overly immoderate measure of a court of law.

Luis A. Ruedas

Thomas More College,
Ohio River Biological Station,
Route 8, Box 86, California,
Kentucky 41007-9804, USA

Tumour diagnosis

SIR — The ability to refine pathological diagnosis and improve its objectivity, particularly in the context of small biopsies and cytological material, has been dramatically demonstrated by the case of Hubert Humphrey's bladder cancer (*Nature* 369, 13; 1994). However, for this approach to become generally applicable in a diagnostic setting will require those tools to be cheap, easy to use, robust and applicable to hospitals and clinics without sophisticated molecular biology laboratories. PCR-based methods are complex, costly and applicable to only relatively few centres, and certainly not in district hospitals at present. Simpler methods exist that are easily and cheaply applicable to clinical samples¹. The approach is based upon the remarkable correlation between over-expression of the p53 protein and malignancy. This has been demonstrated in a wide range of retrospective studies in a wide range of tumour types. In addition, in a large prospective study, p53 immunohistochemistry has been shown to be of real clinical utility in diagnostic cytopathology including in diagnostically difficult cases².

Immunohistochemistry is simple, cheap and employed in nearly all pathology laboratories today. Consequently there is no reason why this simple technology could not be used as an effective adjunct to conventional morphological analysis in cytological and pathological diagnosis and screening today with the attendant positive impact on health care.

Peter A. Hall

Stephanie P. Dowell

David P. Lane

Departments of Pathology and
Biochemistry,
University of Dundee,
Dundee DD1 9SY, UK

1. Hall, P. A. et al. *Lancet* 338, 513 (1991).

2. Dowell, S. P. *Cancer Res.* (in the press).

Prospecting for drugs in ancient texts

Bart K. Holland

The mass screening of plants in the search for new drugs is vastly expensive and inefficient. It would be cheaper and perhaps more productive to re-examine plant remedies described in ancient and mediaeval texts.

THE plants of rainforests are rapidly being destroyed and with them the potential for discovering new drugs based on phytochemicals¹. The huge diversity of species leads to the expectation that many

other compounds are under investigation specifically because of their uses in folk medicine. In Ghana, asthma is treated with a medicinal herb that has recently been shown to contain a compound (dehydrosoyasaponin I, a triterpenoid glycoside) that is the most potent known potassium-channel opener; this may explain the antispasmodic and spasmolytic effects of the plant, which have recently been confirmed *in vitro* and *in vivo*⁴. Traditional Chinese medicinal plants are being investigated as a treatment for eczema⁵, and in the Philippines a tea made from the leaves of *Carmona retusa* is considered to have many medicinal properties; a compound in the leaf is strongly antimutagenic⁶.

There is an extremely large and readily accessible body of traditional medicine, describing a wide range of plants and other substances, that has not recently been investigated systematically. I refer to premodern western medicine, embodied in the writings of ancient Greece and Rome through the Middle Ages and the Renaissance. Many of the reported remedies have been dropped, of course, because they were preposterous⁷. The scientific revolution which began in the mid-sixteenth century dismissed Aristotelian-Galenic medicine as being fraught with errors and as having a stifling influence. This attitude has persisted, causing us to regard ancient and mediaeval medicine as an historical curiosity rather than a source of potential therapies. Images of blood-letting, obnoxious potions and crude, dirty surgery come to mind. But the Greek and Latin herbaria and *materia medica* do contain, in some cases, descriptions of plants of pharmacological effectiveness whose properties have been forgotten.

The plant called silphion, in the genus *Ferula*, was of great economic importance in Greek and Roman times⁸. It had to be gathered and traded; attempts to cultivate it failed, according to Hippocrates. In the second century AD, the gynaecologist Soranus recommended oral administration of silphion sap as a contraceptive; modern experiments show that extracts

from some of silphion's surviving relatives inhibit conception or prevent implantation of a fertilized egg in rodents⁸.

Pennyroyal, an aromatic mint plant (*Mentha pulegium* L.), was considered to be an abortifacient in ancient and mediaeval times. Quintus Serenus wrote in the second century that a fetus less than one month in gestation could be aborted using an infusion of the plant. The active ingredient is pulegone, and recent studies confirm that this chemical induces abortion in animals and people⁹⁻¹² and show that pennyroyal tea would be preferable for this purpose to the purified oil, because the high concentration of pulegone in the oil would be toxic^{13,14}.

Perhaps the time has come to make a relatively small investment in the systematic re-examination of therapies mentioned in Greek and Latin medical texts, through a dialogue between pharmacologists on the one hand, and classicists, mediaevalists and historians of medicine on the other. Such cooperation, which would link ancient texts with modern standards of testing, might result in a useful and inexpensive source of potentially therapeutic compounds. □

Bart K. Holland is in the Department of Preventive Medicine and Community Health, New Jersey Medical School, Newark, New Jersey 07103, USA.

ACKNOWLEDGEMENT. I thank John Riddle for much useful advice.

Dioscorides receives the mandrake root from Euresis, goddess of discovery (AD 512).

therapeutically worthwhile compounds remain undiscovered; but to screen innumerable plants and their compounds in a systematic way is vastly expensive. Random screening is not an effective approach: the National Cancer Institute failed to find a compound with clinical anticancer activity among 114,000 plant extracts from 35,000 species².

The use of folk beliefs and traditional healers as a short-cut to the discovery and isolation of pharmacologically active compounds has been a productive approach. About a quarter of all prescriptions written in the United States are for drugs that contain compounds originally identified from plants², and virtually all currently used drugs derived from plants, including reserpine, quinine, digoxin, digitoxin, tubocurarine, morphine and codeine, were discovered through scientific investigation of folklore claims³. European Renaissance folklore included knowledge of deadly nightshade (*Atropa belladonna*) and mandrake (*Mandragora* species), among many others. Even today, many

1. Fellows, L. *Lancet* **1**, 1330-1333 (1992).
2. Reynolds, T. J. *nath. Cancer Inst.* **83**, 594-596 (1991).
3. Farnsworth, N. R. *Natural Products and Drug Development* (Munksgaard, Copenhagen, 1983).
4. McManus, O. B. *et al. Biochemistry* **32**, 6128 (1993).
5. Atherton, D. J., Sheehan, M. P. & Rustin, M. *Lancet* **II**, 510 (1991).
6. Villaseñor, I. M., Edu, D. A. & Bremner, J. B. *Carcinogenesis* **14**, 123-125 (1993).
7. Burton, J. L. *Lancet* **II**, 1565-1566 (1990).

8. Riddle, J. M. & Estes, J. W. *Am. Scient.* **80**, 226 (1992).
9. Froehlich, O. & Shibamoto, T. J. *agric. Food Chem.* **38**, 2057-2060 (1990).
10. Gordon, W. P. *et al. Drug Metab. Disp.* **15**, 589 (1987).
11. Thomassen, D. *et al. Drug Metab. Disp.* **19**, 997 (1991).
12. Thomassen, D. *et al. J. Pharmac. exp. Ther.* **253**, 567-572 (1990).
13. Gordon, W. P. *et al. Toxic. appl. Pharmac.* **65**, 413-424 (1982).
14. Sullivan, J. B. *et al. J. Am. med. Ass.* **242**, 2873 (1979).



Deadly nightshade (*Atropa belladonna*).

Can the research university survive?

Budgetary and other pressures are changing the character of US universities more quickly than is comfortable, but the process has only just begun. And it is not yet clear where it will end, to judge from a conference last week.

Los Angeles. Is the research university in the United States dead, or dying, or under threat of extinction? And, on the assumption that there is something to salvage from what will otherwise be rubble, what can be done to revitalize the university? That was the theme of a conference organized by the University of California, Los Angeles (UCLA), last week by Dr Kumar Patel, vice-chancellor for research at UCLA and by Dr Roland Schmitt, recently retired as president of the Rensselaer Polytechnic Institute, New York, under the challenging title "Reinventing the research university".

Improvement was more in evidence than re-invention. One speaker, Dr M. R. C. Greenwood, associate director for science at the Office of Science and Technology Policy at the White House, had earlier reminded her audience (university presidents, lesser academic mortals, civil servants and industrialists) of Clark Kerr's dictum, when chancellor of the University of California, that a university is a collection of disparate academic entrepreneurs united only by a "common grievance over parking".

Similarly, the research universities (of which there may now be 200 in the United States) have disparate goals, but are certainly united by their sense of grievance over "indirect costs" and over the federal government's apparent determination that these, the overhead payments accompanying research grants, should be restricted (by formula) and reduced. Luckily, the agenda last week ensured that the issue was only a passing reference in what most speakers said.

But money more generally is the root of the problem. Although the requests in this year's federal budget on behalf of the research-grant agencies include increases greater than the inflation rate, total discretionary expenditure is essentially fixed, so that later this year, and then for many years to come, the US Congress will be able to increase research support only by cutting other spending. (Greenwood said so explicitly; her listeners mostly took the point.) Dr Erich Bloch, president of the US National Science Foundation (NSF) in 1984-90, put it bluntly: the university research system has grown to a size and cost that cannot be sustained under the budget constraints ahead.

There are also other pressures on the system. Governments look to research universities for assurance on matters as different as industrial competitiveness and national security, but demand of the university system as a whole accountability, higher education of better quality and access for

greater proportions of young people. (More than half of high-school graduates in the United States already continue in some form of higher education.) The universities have taken fright at expressions of discontent with the quality of university education from students and their parents (who pay the bills) and employers (or non-employers).

How can the research universities survive these conflicting pressures? It seems agreed that the present is a period of rapid and irreversible change. The golden age of uninhibited growth has ended. But there is no consensus on what happens next. What follows is a thumbnail sketch of several strands of the argument last week.

Improvement, or being better at present tasks, is one tack. The general opinion is that the quality of teaching, especially of undergraduates, needs urgently to be improved. But no research university appears yet to have found a way of rewarding good teachers with promotion, even security. There is also palpable shame at the neglect of curriculum development, but touching faith that new-technology teaching, by means of networks and the like, may be radically more efficient. Self-assessment of effectiveness at reaching declared aims won general approval as a goal, but Schmitt (with practical experience at Rensselaer) declared flatly that most institutions do not even understand how their resources are spent on their different functions.

On the research side, there are several avenues for improvement, but none is decisive. Better management of projects, the more efficient use of what funds there are, more long-term block grants by NSF and the National Institutes of Health (to concentrate particular kinds of research where it is done best), more collaboration between institutions and the pooling of research facilities crop up on different people's menus. In this respect, last week's meeting closely echoed the discussions in Britain over the past decade about arrangements (now in place, but untried) for allocating research funds to the publicly supported universities. The difference was that, last week, the arguments were made by academics, not civil servants or their representatives. Plainly the research universities have seen the writing on the wall.

What do they make of it? For those of nervous disposition, the most alarming presentation would have been that of Dr Alexander MacLachlan, until recently senior vice-president for research at DuPont, who openly explained that his company has abandoned its policy (of the 1970s) of throwing PhDs at

fields of enquiry judged potentially rich in innovation, and instead is concentrating on its core businesses. The spur is the ferocity of international competition.

The corporation insists that its chief expectation of the research universities is a supply of skilled scientists; its recruitment of PhDs should pick up as the recession finally ends, but will not return to the dizzy earlier heights. Another is that DuPont, short of cash from having cut its margins to compete with overseas producers, is now seeking to pool its sense of penury with the penurious universities. MacLachlan said there have already been research contracts with university groups substituting for research that DuPont would previously have done in-house. In Japan, it used to be the case that some university departments would effectively be extensions of a company's research effort; is the United States about to move in that direction while the Japanese government is still formally committed to doubling spending on basic research?

It was plain last week that there are still huge difficulties to be ironed out — the freedom to publish, intellectual property rights and the like. But making university departments or research groups partly responsible for a corporation's research programme is radical enough to count as re-invention.

Dr Don Langenberg, the physicist deputy director of NSF in 1980-82 who is now chancellor of the University of Maryland system, went further with, for example, the concept of "Virtual-U" — a network linking academics with common interests. More radical, for the future of the research university, is the notion that parts of a university might be "unbundled" into a "collection of public enterprises" serving identifiable customers. He offered Maryland's new campus for continuing education as a sign of things to come.

Langenberg is an idealist in looking for an end to bickering between faculty and administration, a realist in acknowledging the importance of customers for university services and a manager tough enough to mould change rather than acquiesce in what happens. That is the sense in which "re-invention" is the right word. Last week's meeting may not have defined what the university of the future will be like, but it powerfully suggested that at least some universities have read the writing on the wall and have determined to make the best of it, however much change will be entailed.

John Maddox

Water and the solid Earth

Charles H. Langmuir

The presence of water is one of the most important attributes of the Earth as a planet, of central importance to the surface envelope of ocean, atmosphere and life. It is also pivotal in the behaviour of the solid Earth — it is incorporated into the ocean crust during hydrothermal circulation around ocean ridge volcanoes, and fluxes the mantle at subduction zones to create the explosive volcanism of convergent margins. Quantitative understanding of the role of water in the solid Earth cycle, however, has been lacking. In particular, the detailed effects of water on mantle melting are poorly known, as are the compositions of hydrous fluids in equilibrium with the mantle. E. Stolper and S. Newman¹ have recently made strides forward on both these subjects.

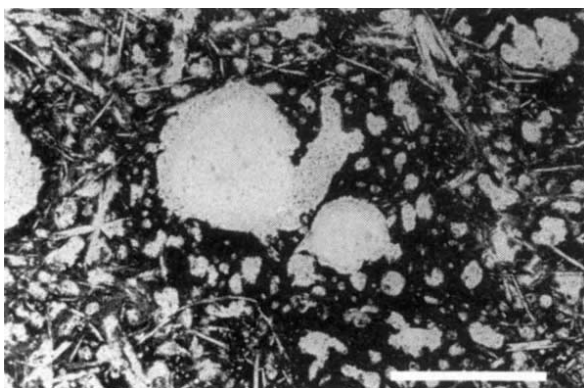
Their progress derives from improved analyses of the water contents of basic magmas. Water has long been recognized as a fundamental agent of magma genesis at convergent margins, and as an important carrier phase for many trace elements, leading to metasomatism and the creation of heterogeneities in the mantle.

Other tracers, such as alkali and alkaline earth elements, B/Be ratios², and uranium-series disequilibria³ have been used successfully to infer the importance of water and the composition of the hydrous phase. But the data on water itself have been sorely lacking, mostly because reliable water analyses require quenched basaltic glasses, and arc volcanics do not include such glasses. Glasses do occur in the back-arc basins behind island arcs, however, and basalts from these regions also have the signature of increased abundances of elements thought to be derived from the slab and carried in a hydrous fluid^{4,5}.

Stolper and Newman have taken advantage of these characteristics of back-arc basin basalts, and determined the water contents of a dozen basalt glasses from the Marianas back-arc basin in the western Pacific. Comparing the H₂O data with other published data from these samples^{6,7}, they show that the volcanic rocks lie on a straight line in composition space. They interpret the data to result from a link between the amount of hydrous fluid added to the source, and the amount the source melts. Because the fluid is rich in some elements, such as Na, Cl, Th, U, Sr and Pb, the addition of the fluid enriches the source in these elements. At the same time, addition of this fluid fluxes the mantle — the more fluid

that is added, the more the melting temperature of the mantle is reduced, leading to a larger extent of partial melting of the mantle.

The idea that a hydrous fluid simultaneously enriches a source in fluid-soluble elements and causes an increase in the extent of melting has been previously suggested to explain data from arc volcanics^{8,9}. The advance of Stolper and



Thin section of a vesicular basalt dredged from the East Scotia Sea back-arc-basin spreading centre (water depth about 3,000 m). The large and abundant vesicles indicate significant amounts of water present in the mantle source of these basalts. Scale bar, 1 mm.

Newman's study is their well constrained calculation for the concentrations of some twenty elements in the hydrous component that has influenced the Marianas back-arc basin. An important and testable aspect of this model is the quantitative estimate of the effect of water (and the elements associated with it) on the extent of melting. They calculate that increasing water from 0.02 to 0.24 weight per cent in the source causes the extent of melting to increase by roughly a factor of four — a huge amount. They then compare the back-arc basalts with those from the neighbouring Marianas arc, and find that the arc basalts are derived by still greater extents of melting, which suggests still higher water contents beneath the volcanic front.

This increase in extent of melting from back-arc to arc has been suggested previously for other back-arc/arc pairs, such as the Izu-Bonin arc and the Bonin back-arc rifts⁵. In island arcs built on thin crust, such as the Marianas, the extent of melting for the arc appears substantially greater than for normal mid-ocean-ridge basalts (MORB), akin to what is observed at the hottest locations on ocean ridges¹⁰.

It is important to recognize that the calculated fluid component is not the fluid that comes off the slab, but rather the composition that has been added to back-arc source regions, which can be several

hundred kilometres from the slab. The fluid flux added to the back-arc source could have undergone a complicated history and be substantially modified from the slab-derived fluid. Stolper and Newman emphasize this point, and assert that the back-arc fluid actually reflects the end product of chromatographic exchange between fluid and mantle. They suggest that the composition they calculate reflects an equilibrium between a fluid and the normal MORB source.

The quantitative constraints on a fluid composition and the inference of how such a fluid influences the extent of melting are well-founded new results of this work, and the arc/back-arc comparison is a useful addition. In combination, these results are a substantial step forward in our understanding of the importance of water to igneous processes.

Stolper and Newman then take the ambitious step of considering some far-ranging implications of their result, including the nature of transport of the hydrous component through the mantle, the systematics of major elements during hydrous melting of the mantle, and the values of partition coefficients between fluid and solid. They suggest that the hydrous component may account for the extent of melting of mantle hotspots, be the primary agent of mantle heterogeneity, and create the characteristic chemical signatures of continents and the MORB source. Of course, many questions arise from the treatment of such a diversity of problems using data from a small number of rocks from a single region.

An important part of the discussion stems from the idea that the calculated fluid is a prototypical hydrous composition in equilibrium with the mantle. On this basis, Stolper and Newman estimate the partition coefficients between solid and fluid (known as *D* values) by dividing their calculated fluid composition by the composition of their normal MORB source. This assumes that the fluid has come to equilibrium with normal MORB mantle, and no longer carries the slab signature. A concern is that the *D* values are very low, and they imply very high fluid solubility for almost all incompatible elements, in apparent contradiction of experimental results¹¹.

More important, elements considered to be insoluble in fluids appear highly enriched in the calculated fluid: Zr has the same *D* as Na; Th is calculated to be similar to U, and not much different from Cl. At the same time, the calculation assumes that Ti, another supposedly insoluble element, is absent from the fluid component. These calculated results will no doubt soon be subject to experimental

A. D. Saunders

tests, as will the relationship between water content and extent of melting.

Some will also question the pertinence of the results to the origin of continents. It has long been known that continents and the normal MORB source are complementary, and that the difference can be modelled approximately for trace elements by the extraction of a small-degree melt from the mantle. Stolper and Newman's fluid is very like a small-degree melt, and the uncertainties on their inferred *D* values are sufficiently high that it is not obvious that a hydrous fluid and a small-degree melt can be distinguished. Furthermore, key ratios that are different in continental and oceanic rocks (such as Ce/Pb) are not fractionated sufficiently by the fluid, while others (such as Sr/Nd) are fractionated too much.

Other important viewpoints will come from studies of other back-arc basins, and further quantification of fluids associated with arc and back-arc volcanism. Studies of arc volcanoes suggest a rather different fluid behaviour than that which Stolper and Newman infer: elements such as B, Cs, Ba, Sr and Pb appear mobile in fluids, whereas elements such as Zr and Th do not^{8,9}. The chemistry of back-arc basin basalts varies from one basin to another, and some data from other back-arc basins do not fall on the Marianas trends. Therefore it is probable that inferred fluid compositions for other back-arc basins will differ in interesting ways. Quantification of chemical components from samples from many different settings, combined with experimental data on partitioning between fluid and solid and on hydrous mantle melting, should lead to rapid further progress.

Although many questions remain, the fact that Stolper and Newman can address such a variety of problems underlines the importance of their results, and the essential role of water in the plate-tectonic geochemical cycle. □

Charles H. Langmuir is at the Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA.

Parasites and polymorphisms

Dieter Ebert and Roberto Lorenzi

It was Haldane's suggestion¹ that there is an evolutionary arms race between hosts and pathogens, one characterized by the host's need for continual adaptation to the coevolving antagonist. In this light, genetic recombination could be advantageous because parasites are believed to adapt primarily to common host genotypes. This so-called Red Queen hypothesis provides a powerful, yet not fully understood, mechanism for the maintenance of genetic polymorphism as found in both plants and animals. The part played by genetic diversity in the coevolution of hosts and parasites was the subject of two interdisciplinary meetings last month*, bringing results together from medicine and theoretical, organismic and molecular biology.

One of the best examples of host evolution in response to a pathogen comes from Africa, where infectious diseases continue to be a major cause of death in human populations. In vertebrates, the highly polymorphic glycoproteins coded by the major histocompatibility complex (MHC) bind foreign peptides in a sequence-dependent fashion, and thereby induce a T-cell immune response against specific pathogens. Evidence for parasite-driven changes in the allelic frequency of MHC genes came from the finding that resistance to *Plasmodium falciparum* malaria in West Africa is associated with a human class I allele, HLA-B53, and an unusual class II haplotype. Both genes are under-represented in children suffering from severe malaria, and are more common in the overall population than elsewhere in the world, suggesting that selection by malaria has contributed to the increase in frequency of these protective alleles in West Africa². This is in agreement with time-lagged, frequency-dependent coevolution, because it might take many generations for the parasite to evolve an evasion strategy after the host has evolved resistance.

Complexity

The finding that protection against malaria in East Africa operates with different MHC haplotype associations from those in West Africa highlights the complexity of parasite selection (A. V. S. Hill, John Radcliffe Hospital, Oxford). In particular, it is intriguing that this difference is apparently not due to B53-epitope variation between East and West African strains of *P. falciparum*. The explanation for this spatial variation is probably com-

plicated, as differential selection on MHC genes over time and space is likely to be driven not only by malaria but also by other diseases. The fact that we detect an MHC-malaria association at all underlines the selective strength of malaria in driving host-parasite coevolution.

Fitness

Differential fitness of host genotypes, as found in the malaria example, might result in a frequent loss of obsolete resistance genes. But because the strong polymorphisms of the MHC and of disease-resistance genes in plants appear to be ancient, a storage mechanism seems to operate to prevent the loss of alleles. The polygene storage concept of Mather³ suggests that genes can be maintained by keeping them in close linkage, as is the case in both vertebrate (J. Klein, Max Planck Institute for Biology, Tübingen) and plant resistance systems (I. R. Crute, Horticulture Research International, Warwick). Low recombination rates will preserve alleles from extinction, but release them from time to time, and once favourable they might increase again⁴. So during sexual recombination new alleles and gene combinations arise; combinations of linked genes are maintained to some degree, however, facilitating preservation of currently obsolete genes.

In contrast to host evolution, pathogen evolution has been observed since early attempts to select for resistance in crops. But the evolutionary mechanisms involved, and the extent of genetic diversity of parasites, is only beginning to emerge. In Papua New Guinea, polymorphism of merozoite surface antigens of *P. falciparum* was evident in 68 per cent of the infected human population (K. Day, Univ. Oxford). This is probably a result of repeated infections — in areas where malaria is endemic, within-host diversity is correlated with transmission probability, resulting in a variable degree of inbreeding of the sexual state of *P. falciparum*. Similarly, within-host diversity of a trypanosome parasite that infects bumble bees has been found to be unexpectedly high, which is also probably the result of repeated infections (P. Schmid-Hempel, ETH-Zürich).

High frequency of multiple infection is believed to play an important role in parasite evolution⁵. Whereas between-host competition of parasite strains selects for the highest transmissibility, within-host competition selects for survival and reproduction of the parasite. For example, if parasites outcompete each other through rapid reproduction, and if reproduction is correlated with virulence,

1. Stolper, E. & Newman, S. *Earth planet. Sci. Lett.* **121**, 293–325 (1994).
2. Morris, J., Leeman, W. & Tera, F. *Nature* **344**, 31–36 (1990).
3. Gill, J. B. & Williams, R. *Geochim. cosmochim. Acta* **54**, 1427–1442 (1990).
4. Saunders, A. D. & Tarney, J. *Geochim. cosmochim. Acta* **43**, 555–572 (1979).
5. Fryer, P., Taylor, B., Langmuir, C. & Hochstaedter, F. *Earth planet. Sci. Lett.* **100**, 161–178 (1990).
6. Hawkins, J., Lonsdale, P., Macdougall, J. & Volpe, A. *Earth planet. Sci. Lett.* **100**, 226–250 (1990).
7. Volpe, A. et al. *Earth planet. Sci. Lett.* **100**, 251–264 (1990).
8. Miller, D., Langmuir, C., Goldstein, S. & Franks, A. *J. geophys. Res.* **97**, 321–343 (1992).
9. Lühr, J. J. *Volcan. geotherm. Res.* **54**, 1–18 (1992).
10. Plank, T. & Langmuir, C. *Earth planet. Sci. Lett.* **90**, 349–370 (1988).
11. Tatsumi, Y., Hamilton, D. & Nesbitt, R. J. *Volcan. geotherm. Res.* **29**, 293–309 (1986).

* Infection, Polymorphism and Evolution, Royal Society, London, UK, 25–26 May 1994. Adaptations for Disease Resistance and Virulence, Ciba Foundation, London, UK, 27 May 1994.

within-host selection for faster reproducing strains will increase parasite virulence and maintain genetic diversity, at the cost of reduced transmissibility (M. Nowak, Univ. Oxford).

To increase survival within hosts, some viruses, such as the human immunodeficiency virus, have evolved high mutation rates. The consequence is the generation of rare genotypes which have temporary advantages in the frequency-dependent race with the host immune response (A. Leigh Brown, Univ. Edinburgh). This high mutation strategy may be disadvantageous to both host and parasites if the within-host competition leads to increasingly higher virulence and rapid death of the host⁶. But it can also generate synergistic evasion strategies of the host immune system, which require epitope variability to operate⁷.

High rates of generic change — through mutations, sexual recombination or both — characterize host-parasite co-

evolution. Genetic diversity provides the ammunition in this arms race, in which parasites evolve in a constant conflict of within and between-host selection, and hosts are under constant pressure to escape their parasites. Whether or not sex is crucial in this race remains one of the most exciting questions in biology. □

Dieter Ebert is at the Centre for Population Biology, Imperial College at Silwood Park, Ascot, Berks SL5 7PY, UK. Roberto Lorenzi is in the Department of Immunology, The Babraham Institute, Babraham, Cambridge CB2 4AT, UK.

1. Haldane, J. B. S. *La Ricerca Scientifica* **19** suppl. 19, 68–75 (1949).
2. Hill, A. V. S. *et al. Nature* **352**, 595–600 (1991).
3. Mather, K. *Symp. Soc. exp. Biol.* **7**, 66–95 (1953).
4. Hamilton, W. D. J. *Hered.* **84**, 328–338 (1993).
5. Herre, E. A. *Science* **259**, 1442–1445 (1993).
6. Levin, B. R. & Bull, J. J. *Trends Microbiol.* **2**, 76–81 (1994).
7. Allen, P. M. & Zinkernagel, R. M. *Nature* **369**, 355–356 (1994).

INTERPLANETARY DUST

The ring around us

Donald E. Brownlee

ONE of the Solar System's great failures was its inability to assemble a full-scale terrestrial planet between Mars and Jupiter. The orderly fabrication schedule of what would have been the fifth terrestrial planet was rudely interrupted by the rapid and nearby formation of Jupiter. Today, all that remains of this process is the asteroid belt, a collection of relict planetary building blocks up to 1,000 km in diameter.

Following an initial period of growth, the main-belt asteroids have endured aeons of destructive collisions that slowly but relentlessly pulverize them. As dust particles are generated they migrate towards the Sun, forming a thin particle sheet whose radial distribution can be strongly influenced by the gravitational pull of the planets. On page 719 of this issue, Dermott *et al.*¹ suggest that perturbations by the Earth stall a portion of the inflow to produce a circumsolar dust ring centred just beyond the Earth's orbit. The ring has peculiar azimuthal fine structure, with a gap that the Earth resides in and a dust concentration that trails the Earth like a shadow as it orbits the Sun.

The origin, evolution and destruction of dust in our Solar System and in other circumstellar dust systems are a complex set of processes². Interplanetary dust originates from comet sublimation, asteroid collisions and direct injection of interstellar dust from the Galaxy³. It is removed by ejection from the Solar System, collision with other particles, accretion onto planets or evaporation near the Sun. The

timescales for these processes are short, and individual 10- μ m particles survive in the inner Solar System only for about 10⁴ years.

One of the key cleansing features in the Solar System is the Poynting–Robertson (P–R) effect⁴, the orbital decay of small particles due to the drag component of sunlight pressure. This is a universal process that occurs for all orbiting particulates where absorbed and scattered photons yield a net drag force against a particle's motion. For a 10- μ m particle, this effect produces an inwards spiral of about one Earth radius per year at a distance of 1 AU from the Sun. The effect drives particles towards the Sun where they are vaporized or fragmented by collisions, in which case their debris may be driven out of the Solar System by radiation pressure as beta meteoroids⁵. As particles spiral inwards from the asteroid belt they can be temporarily trapped in resonances that stop orbital decay⁶. Simulations have shown that particles of the appropriate size can be trapped for times as long as 100,000 years, although this is a highly stochastic process and many particles escape without significant trapping.

Resonance trapping by the Earth occurs when a particle orbital decays to a location where its orbital period is a simple fraction (6/5, 7/6 and so on) of a year. Quasi-stable trapping occurs at these locations because close resonant encounters with the Earth add sufficient orbital energy to a particle to counteract the constant energy loss from P–R drag. As viewed from a Sun-

centred reference frame that co-rotates with the Earth, the path of a trapped dust particle is an epicycloid with multiple loops approaching the Earth's orbit at the particle's closest approach to the Sun.

Orbital calculations by Jackson and Zook⁶ suggested that the stalling of asteroid dust in the many resonances just outside the Earth's orbit could result in a dust ring. Numerical simulations by Dermott *et al.* back this up and, in addition, they show that structure in the ring leads to an asymmetry in the concentration of dust leading and trailing the Earth in its orbit about the Sun. Resonances with different orbit-period commensurabilities all have the common property that one of their epicycle loops occurs just behind the Earth, producing a dust concentration there. Just in front of the Earth there is a gap in the ring, but elsewhere the ring is relatively uniform because epicycle loops of various resonances are out of phase. The predicted variations of dust density just leading and trailing the Earth are quantitatively consistent with asymmetry seen in the infrared brightness of the zodiacal light in these directions as measured by the IRAS spacecraft.

One of the interesting predictions of this model is that transfer of trapped asteroidal particles to the Earth will have annual variations. Because of its orbital eccentricity, the Earth drifts within the gap in the ring and near the end of each year it lies closest to the trailing dust cloud where enhanced accretion of asteroid dust can occur. This annual enhancement of asteroid dust should show up in the micrometeorite collections routinely made in the stratosphere. One of the goals of collecting stratospheric micrometeorites is the identification of cometary and asteroidal materials, and this yearly variation may provide an important criterion for distinguishing between particle types. Rietmeijer⁷ has already found evidence for temporal changes in particle types that might be due to such annual variations. And finally, looking further afield, rings of resonantly trapped dust undergoing P–R drag might be a clue to otherwise undetectable planets in the dusty disks that commonly surround other stars⁸. □

Donald E. Brownlee is in the Department of Astronomy, University of Washington, Seattle, Washington 98195, USA.

1. Dermott, S. F., Jayaraman, S., Xu, Y. L., Gustafson, B. Å. S. & Liou, J. C. *Nature* **369**, 719–723 (1994).
2. Grün, E. *et al. Icarus* **62**, 244–272 (1985).
3. Grün, E. *et al. Nature* **362**, 428–430 (1993).
4. Burns, J. A., Lamy, P. L. & Soter, S. *Icarus* **40**, 1–48 (1979).
5. Zook, H. A. & Berg, O. E. *Planet. Space Sci.* **23**, 183–203 (1975).
6. Jackson, A. A. & Zook, H. A. *Nature* **337**, 629–631 (1989).
7. Rietmeijer, F. in *Analysis of Interplanetary Dust* (eds Zolensky, M., Rietmeijer, F. & Zook, H.) 255–276 (Am. Inst. Physics, New York, 1994).
8. Roques, F. *et al. Icarus* **108**, 37–58 (1994).

Glia and neurons in dialogue

David Attwell

NEURONS are the computational elements in the brain, their function being to modify signals and transmit them to other neurons. The numerically superior glial cells merely support neuronal function by controlling the concentrations of neurotransmitters and ions outside the neurons. In a nutshell, these are the tenets that dominate current views of information processing in the brain.

Two papers, one by Parpura *et al.* on page 744 of this issue¹ and one by Nedergaard in *Science*², challenge this dogma by showing that rises of the calcium concentration, $[Ca^{2+}]_i$, in glial cells (astrocytes) can cause $[Ca^{2+}]_i$ rises in surrounding neurons. The implication is that glia may well be involved in information processing. But although both papers report the same basic observation, they provide contradictory data on the mechanisms by which calcium signals pass from glia to neurons.

Parpura and colleagues report that a $[Ca^{2+}]_i$ rise in cultured astrocytes, produced with bradykinin, mechanical stimulation or strong illumination, leads both to the release of the excitatory neurotransmitter glutamate from the astrocytes

and to a rise in $[Ca^{2+}]_i$ in adjacent neurons. The rise in neuronal $[Ca^{2+}]_i$ could be blocked with the glutamate analogue D-2-amino-5-phosphonopentanoic acid, which suggests that it is generated by the released glutamate activating N-methyl-D-aspartate (NMDA) receptor channels — a type of glutamate-gated membrane ion channel which is known to have a high permeability to calcium.

Conventionally it is neurons that release glutamate in the nervous system. How might astrocytes do this? There are three well-known mechanisms of glutamate release: calcium-dependent exocytosis of glutamate stored in vesicles³; reversed operation of sodium-dependent uptake carriers⁴, which normally remove glutamate from the extracellular space to terminate synaptic transmission; and release through a furosemide-sensitive anion channel or carrier which can be activated by astrocyte swelling⁵. Despite its calcium dependence, the glutamate release reported by Parpura *et al.* is unlikely to be by exocytosis; this is because the astrocytes studied expressed no synaptotagmin, a vesicular protein involved in calcium-dependent exocytosis³.

Glutamate release is also unlikely to be by reversed uptake, because this process does not require a rise in $[Ca^{2+}]_i$ and because blockers of the operation of glutamate-uptake carriers did not affect glutamate release from astrocytes¹.

Finally, although Parpura *et al.* found that the bradykinin-induced glutamate release from astrocytes was blocked by furosemide, it was also blocked by removing external calcium; yet the swelling-activated, furosemide-sensitive mechanism of glutamate release studied previously⁵ is unaffected by removal of calcium⁶. So unless a rise in $[Ca^{2+}]_i$ and cell swelling can independently activate glutamate release by a common mechanism, the mode of release in the latest experiments¹ remains unclear.

In contrast to the results of Parpura *et al.*, Nedergaard found that the rise in neuronal $[Ca^{2+}]_i$ evoked by a rise in glial $[Ca^{2+}]_i$ was unaffected by blocking NMDA receptors and by calcium removal, but was inhibited by agents that block gap junctions. Nedergaard therefore attributed the $[Ca^{2+}]_i$ signal to calcium release from intracellular stores, and postulated that it spread through gap junctions between astrocytes and neurons. There is no easy way to reconcile the two sets of data^{1,2}; perhaps there are two independent mechanisms for transmission of calcium signals from glia to neurons.

VISION

Driven around the bend in Scotland



VEHICLE manufacturers around the world are investing a great deal in attempts to develop intelligent guidance systems. But how do human drivers use visual information from the road ahead to steer their cars? On page 742 of this issue M. F. Land and D. N. Lee address one aspect of the question — that of where drivers fixate their gaze. Using an instrumented car equipped with a computerized video monitor, the authors recorded the directions of eye and head movements as drivers negotiated a winding road. By choosing a one-way street (around Arthur's Seat, a large hill in the centre of Edinburgh), they eliminated the problem of oncoming traffic. Once stripped to its essentials, the answer turned out to be surprisingly simple: as they come around the curves, drivers show a strong tendency to fixate along a tangent from the car to the inside edge of the road (picture on the left, above). The direction of



gaze is highly correlated with the direction of steering, and the authors suggest that the angle between the line of gaze and the direction of movement may be used as a simple predictor of the sharpness of the bend ahead, and hence of how the steering wheel should be positioned. Of course, human drivers are sometimes distracted (in the instance shown here, by a jogger of the opposite sex: right, above), and some are more easily distracted than others (see Fig. 1 of the paper). Nevertheless, their gaze consistently returns to the edge of the road. Besides offering the prospect of a somewhat Orwellian way of investigating our inner thoughts, these results will no doubt stimulate new approaches to the design of intelligent vehicles. It seems likely that they will also prompt further investigation into how people select visual information during the performance of other complex tasks.

Charles Jennings

What is the functional significance of the newly discovered signalling pathway? Increases of glial cell $[Ca^{2+}]_i$ can be provoked by the opening of ion channels by fast transmitters such as glutamate^{7,8} and subsequent entry of calcium into the cell, as well as by calcium release from intracellular stores following stimulation of phospholipase C by agents such as bradykinin, noradrenaline, ATP and glutamate⁸. The resulting rise of $[Ca^{2+}]_i$ in neurons may modulate the electrical activity of neurons by activating calcium-gated ion channels directly, or by affecting intracellular messenger systems controlling the phosphorylation of ion channels. Such an interaction could, in principle, affect signal processing both locally and at a distance. First, it could amplify local rises of $[Ca^{2+}]_i$ in neurons: for example, glutamate released from neurons as a result of a rise in neuronal $[Ca^{2+}]_i$ may raise the $[Ca^{2+}]_i$ in glia, which then produce a further increase in the neuronal $[Ca^{2+}]_i$. Second, long-distance propagation of $[Ca^{2+}]_i$ waves through astrocytes coupled by gap junctions⁹ might lead to a rise of $[Ca^{2+}]_i$ in neurons far from the location at which the glial $[Ca^{2+}]_i$ was initially elevated.

The possibility of signals passing from neurons to glial cells and thus to other neurons is interesting. It will doubtless be invoked to explain signal transmission currently attributed to unidentified diffusible extracellular messengers, for example during the neuronal depolarization that occurs in spreading depression (discussed by Nedergaard), and during the storage of information as an altered synaptic strength in the model of memory called long-term potentiation¹⁰. Speculation on this second possibility should be tempered by the view¹¹ that "the only feature which can really be said to be shared by memory and glia is that very little is known about either of them". That statement, however, now has less force than previously, for the results of Parpura *et al.* and Nedergaard significantly reduce our ignorance of glial cells. □

David Attwell is in the Department of Physiology, University College London, Gower Street, London WC1E 6BT, UK.

Flowers and funerals

Peter D. Moore

SAYING it with flowers is not a modern idea. Quite apart from the extensive use of flowers as symbolic items in art over many centuries, there is also evidence of floral tributes being used in graves since prehistoric times. The detective work involved is by no means straightforward. But it can provide pointers to prehistoric ritual and habits, as Richard Tipping shows in the *Journal of Archaeological Science*¹.

In the New Kingdom of Egypt, the tomb of Tutankhamun himself contained



Scented tribute — meadowsweet (*F. ulmaria*).

an abundance of floral garlands², and flowers have even been found associated with Neanderthal burials in Iraq³. Further north, examples are much rarer: one of the few situations in which intact plants have been found in Bronze Age (2000–500 BC) burials of northern Europe is the oak coffin burial of a girl at Egtved, Jutland⁴, in which flowers of yarrow (*Achillea millefolium*) and fronds of bracken (*Pteridium aquilinum*) were discovered. In Scotland, the use of floral tributes in Bronze Age graves has long been suspected. The evidence has been controversial, however, and Tipping has used pollen analysis to put the case on a firmer footing.

The real problem is one of preservation. In Tutankhamun's tomb, flowers of cornflower (*Centaurea depressa*) and the blue water lily (*Nymphaea caerulea*) were superbly preserved in the extremely dry conditions of the burial, but were so delicate that they often fell to dust as soon as they were touched. In Denmark, the anaerobic conditions within the oak coffin of the Egtved burial undoubtedly contri-

buted to the excellent condition of corpse, clothing and the botanical material. But such conditions are exceptionally rare, and flowers are particularly delicate plant organs that seldom survive any form of fossilization. This is not so for pollen grains, however, which have a robust outer coat that often survives in archaeological sites, both in waterlogged and in very dry conditions.

The application of pollen analysis to the investigation of Scottish Bronze Age burials is not new. In 1978 Dickson⁵ analysed black, crumbly material from a barrow at Ashgrove, Fife, and produced surprising results, for the material contained an abundance (54%) of lime tree (*Tilia*) pollen and a significant amount (15%) of meadowsweet (*Filipendula*) pollen. The lime tree is not regarded as native to Scotland, so Dickson came to the conclusion that the material had been derived from a drink of imported mead that had spilled from an overturned beaker within the burial, the honey used in the mead having been derived from the nectar of lime. This proposal also explained the presence of the meadowsweet, which has traditionally been used as an additive to mead⁶. The inclusion of mead drinks in burials had not been previously recorded in Scotland, but was known from Bronze Age Denmark, including the oak coffin at Egtved.

Tipping has now analysed the pollen at four other Scottish burials (three sites from Perthshire in eastern Scotland and one on the island of Orkney in the north). Like the Ashgrove site, the three Perthshire burials also had traces of a dark stain in the sand below the decayed bone fragments of the corpse (at Ashgrove, the stain was thought to be traces of the spilled mead). The Orkney site contained both inhumations and an urn with cremated bone dust and some plant fibres. Pollen analysis from all four of the sites revealed very high proportions of *Filipendula* (up to 40%), but virtually no *Tilia* or any other likely pollen component of honey. So the mead hypothesis, though still a possible explanation for the Ashgrove burial, is not tenable at these sites.

Such high proportions of *Filipendula*

Erratum

In the 21 April issue of *Nature*, A. L. Greer's News and Views article incorrectly gave the address of R. Nagarajan and K. Chattopadhyay as Bangalore University rather than the Indian Institute of Science, Bangalore. In the same article, reference should have been made to the work of H. Chen, Y. He, G. J. Shiflet and S. J. Poon (*Scripta metall. mater.* **25**, 1421–1424; 1991).

1. Parpura, V. *et al.* *Nature* **369**, 744–747 (1994).
2. Nedergaard, M. *Science* **263**, 1768–1771 (1994).
3. Jessell, T. M. & Kandel, E. R. *Neuron* **10** Suppl. 1–30 (1993).
4. Szatkowski, M., Barbour, B. & Attwell, D. *Nature* **348**, 443–446 (1990).
5. Kimelberg, H. K. *et al.* *J. Neurosci.* **10**, 1583–1591 (1990).
6. O'Connor, E. R. & Kimelberg, H. K. *J. Neurosci.* **13**, 2638–2650 (1993).
7. Müller, T. *et al.* *Science* **256**, 1563–1566 (1992).
8. Cornell-Bell, A. H. & Finkbeiner, S. M. *Cell Calcium* **12**, 185–204 (1991).
9. Cornell-Bell, A. H., Finkbeiner, S. M., Cooper, M. S. & Smith, S. J. *Science* **247**, 470–473 (1990).
10. Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
11. Katz, B. *Nerve, Muscle and Synapse 7* (McGraw-Hill, New York, 1966).

pollen are quite exceptional for any recorded soil sample that has been submitted to pollen analysis. In his report of analyses of Funnel-Beaker (middle Neolithic, around 3000 BC) burials in Denmark, for example, Andersen⁷ records an abundance of lime (in this case probably a dominant local tree), but no *Filipendula*. It is extremely unlikely that the consistently high values for this insect-pollinated taxon in the Scottish burials simply reflect a natural background pollen fallout.

The genus *Filipendula* contains two British species that are indistinguishable on the basis of their pollen grains. Of the two, however, the meadowsweet (*F. ulmaria*) is by far the most likely species to be represented in burials because it is more common, is more robust as a plant, forms denser populations (and is therefore more easily gathered) and is not confined to chalk and limestone as is the dropwort (*F. vulgaris*). Meadowsweet produces a profuse abundance of creamy-white flowers that emit a strong, sickly-sweet odour, and this, as Tipping claims, could provide the most obvious explanation

for its high representation in these Scottish burials; a mattress of meadow-sweet flowers would not only look attractive, but would mask any odour of corruption.

Whether the plant was credited with any further magical properties of preservation or raising the spirits (as in mead) can only be a matter for conjecture, but the Bronze Age people of Scotland do seem to have been early in their expression of grief in floral terms. Although they seem to have been prepared to import such essentials as mead, the Bronze Age Scots were clearly more economical in their attitude to luxuries like floral tributes and used home-grown materials. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

1. Tipping, R. J. *archaeol. Sci.* **21**, 133–139 (1994).
2. Hepper, F. N. *Pharaoh's Flowers* (HMSO, London, 1990).
3. Leroi-Gourhan, A. *Science* **190**, 562–564 (1975).
4. Glob, P. V. *The Mound People* (Faber & Faber, London, 1970).
5. Dickson, J. H. *Antiquity* **52**, 108–113 (1978).
6. Moore, P. D. *Nature* **276**, 443 (1978).
7. Andersen, S. T. J. *Eur. Archaeol.* **1**, 153–180 (1990).

CLIMATE CHANGE

Outlook becoming hazier

Tom M. L. Wigley

CAN particles so small that millions could dance on the head of a pin really affect the climate of the globe? The answer appears to be a resounding 'yes', as shown by K. E. Taylor and J. E. Penner on page 734 of this issue¹. Aerosols, abundant in the air we breathe, come from many sources, land and ocean, have a variety of chemical compositions, and span a size range from submicroscopic to almost visible. Although we cannot see the individual aerosol particles, we can see their effects — they are the main cause of the haze that reduces visibility over wide areas of industrialized Europe and North America.

The climatologically important aerosols range from 0.1 µm to 1.0 µm in diameter, and the most important of these are sulphate compounds — droplets of sulphuric acid and ammonium sulphate. Centuries ago, when the air was much cleaner, most of these sulphate aerosols came from the ocean. A biological waste product, dimethyl sulphide (or DMS), emitted into the atmosphere from the ocean surface, was the primary source; DMS and its products are oxidized in the atmosphere and in clouds by powerful natural oxidizing agents to form sulphate aerosols. Today, the main source is industrial activity, outstripping the natural 'background' source by a factor of three to one in global mean terms. Specifically, the progenitor of most of today's sulphate

aerosols is sulphur dioxide released at the rate of some 70–80 Tg sulphur per year (about 5 tonnes of SO₂ per second), mainly from fossil fuel combustion.

The cooling effect of man-made aerosols and their role in climate change was first suspected many years ago. The eminent climatologist J. Murray Mitchell Jr, who died a few years back, speculated in the 1970s on the possible compensating effects of aerosols and greenhouse gases (see, for example, ref. 2). But early attempts to quantify the aerosol term seemed to indicate that it was small, and the focus of interest shifted to the greenhouse effect.

In recent years, the pendulum has swung the other way, largely through the work of Bob Charlson and colleagues. In 1987 (ref. 3) they highlighted work by Twomey⁴ on what is now referred to as the indirect aerosol effect. Because sulphate aerosols can act as nuclei on which cloud droplets condense, more aerosols mean more droplets. For fixed availability of water, this means that the average droplet radius in a cloud must be smaller; and smaller droplets are more reflective. The net consequence of a greater aerosol concentration, therefore, is higher cloud albedo or reflectivity, more incoming solar radiation reflected back into space (a 'negative forcing'), and hence a cooling effect.

In reality, these processes are complicated and it is still not possible to quantify this indirect effect reliably. More easily quantified, but still highly uncertain, is the so-called direct aerosol effect, in which the aerosols themselves act as the reflecting agents, under clear sky conditions, and the net effect of more aerosols is also cooling⁵. It is the climate effect of this process that Taylor and Penner have examined.

Averaged worldwide, aerosol cooling may partly offset greenhouse warming, but this is a gross oversimplification of the problem. Because their lifetime is only a few days, the main concentrations of fossil-fuel-derived sulphate aerosols are near their sources, so the cooling influence should occur mainly in the Northern Hemisphere close to the industrialized regions of North America and Europe. The patterns of direct aerosol forcing, first derived by Charlson *et al.*⁵, do indeed show this very uneven spatial distribution. But because of the motion of the atmosphere and feedback effects associated with, for example, sea-ice changes, the climate response is most unlikely to mirror the forcing pattern. Taylor and Penner show just how different the forcing and response patterns may be. Although the cooling pattern is still highly variable spatially, it is not confined to the main forcing regions, nor just to the Northern Hemisphere. Significant cooling occurs over the whole globe.

More important than the aerosol-alone or CO₂-alone response patterns is the temperature response to combined CO₂ and aerosol forcing. With CO₂ alone, thought to be the dominant anthropogenic climate influence until recently, one would expect warming worldwide, probably amplified in high latitudes. When aerosols are added, the picture is entirely different — there are both regions of cooling and regions of warming. The quantitative details of Taylor and Penner's results may be subject to revision, but qualitatively they are almost certainly correct; and this will radically change our idea of how human activities might have affected the climate. If searching for the fingerprint of anthropogenic climate change⁶ is like searching for a needle in a haystack, then we've been looking in the wrong haystack.

The consequences for the future are both more important and harder to deal with. Judging from scenarios developed for the Intergovernmental Panel on Climate Change⁷, aerosol forcing (that is, the change in radiative flux reaching the Earth) due to changes in emissions of SO₂ from fossil-fuel combustion is a significant fraction of the total anthropogenic forcing worldwide⁸. Spatially, aerosol forcing is much more important, and the spatial details of future changes in climate may well be dominated by the effects of

changes in the patterns of SO_2 emissions. In fact, even in forecasts for which the global-mean change in SO_2 emissions is minimal, a change in the source distribution, say from a European and North American focus to one centred on eastern Asia, could alter the pattern of climate change considerably.

So far, all spatially specific projections of future climate change^{9,10} have assumed CO_2 to be the only forcing agent. Because the models used still have recognized deficiencies and because their predicted changes differ widely at the regional level, the value of these projections could be questioned even if CO_2 were the only forcing agent. The fact that aerosols have been ignored means that projections may well be grossly in error over much of the world. Not only are the magnitudes of temperature change likely to be overestimated almost everywhere, but, unless the effects of aerosols are included, even the sign of the estimated changes may be wrong in some areas. Furthermore, aerosols are virtually certain to affect all climate variables, not just temperature. This adds another straw to the camel's back for climate modellers.

That challenge is multiplied for those who seek to use the results of these models. With CO_2 alone, it is relatively easy to make use of a single experimental result, such as that obtained by doubling the amount of CO_2 in the model, to estimate changes for a wide range of situations. With aerosols included, each pattern for future emissions will produce its own unique climate-change signature. To cover the range of possible emission patterns may require many, many modelling experiments. Perhaps the biggest challenge, therefore, is to devise and apply a subset of emission patterns which might be linearly combined to span a much wider range of possibilities.

But is this linear superposition approach valid? Nonlinearities may manifest themselves in two ways, at the global-mean level and in the patterns of climate change. For the first, the key parameter is the climate sensitivity, defined as the equilibrium global-mean temperature change per unit change in radiative forcing. This parameter, λ , uncertain by a factor of at least three, is thought to lie between 0.3°C and 1.0°C for every 1 W m^{-2} of forcing⁹ with about 90 per cent confidence. But Taylor and Penner's results indicate that the climate sensitivity for CO_2 -induced forcing may differ from that for aerosol-induced forcing — in other words, the sensitivity may depend on the type and pattern of forcing. If correct, this is a startling and somewhat puzzling result of considerable importance.

For CO_2 , their model gives a value of $1.7^\circ\text{C W}^{-1}\text{ m}^2$, an unusually high value. For aerosol forcing, λ is substantially less,

$1.0^\circ\text{C W}^{-1}\text{ m}^2$. It is not the absolute values here that are important, but the relative values. These add a new element of uncertainty to predictions of future human-induced climate change. Apart from complicating the concept of climate sensitivity that lies at the heart of many climate analyses, it means that estimating future global-mean temperature changes from simple energy-balance models that combine radiative forcings from different sources (as in ref. 8, for example) may not be valid.

Notwithstanding the climate sensitivity problem, one can ask whether the pattern of change for combined CO_2 and aerosol forcing is the same as the sum of the patterns for CO_2 alone and aerosols alone. Reassuringly, it appears that the patterns are similar — the spatial correlation between them for annual temperatures is 0.86 (K. E. Taylor, J. E. Penner and B. D. Santer, personal communication). So it may still be possible to estimate climate change for future combinations of greenhouse gas and SO_2 emissions by combining results from simpler experiments, although the assessment will undeniably be complicated by the differential sensitivity effect.

Clearly, it is vital to confirm Taylor and Penner's results with other models. If they are correct, however, their importance should not be underestimated. They imply that we are still on the upward slope of the climate-change learning curve and that it may take longer than we thought to reduce uncertainties in climate predictions. In more ways than one, therefore, aerosols add a new dimension to the problem of predicting climate change and providing results that are of use to policy makers. The outlook may indeed be warmer, but our picture of it has become much hazier. □

Tom M. L. Wigley is at the University Corporation for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA.

1. Taylor, K. E. & Penner, J. E. *Nature* **369**, 734–737 (1994).
2. Mitchell, J. M. Jr in *The Changing Global Environment* (ed. Singer, S. F.) 149–173 (Reidel, Dordrecht, 1975).
3. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
4. Twomey, S. A., Peipgrass, M. & Wolfe, T. L. *Tellus* **B36**, 356–366 (1984).
5. Charlson, R. J., Langner, J., Rodhe, H., Leovy, C. B. & Warren, S. G. *Tellus* **AB43**, 152–163 (1991).
6. Santer, B. D., Wigley, T. M. L. & Jones, P. D. *Clim. Dynam.* **8**, 265–276 (1993).
7. Leggett, J. A., Pepper, W. J. & Swart, R. J. in *Climate Change 1992. The Supplementary Report to the IPCC Scientific Assessment* (eds Houghton, J. T., Callander, B. A. & Varney, S. K.) 69–95 (Cambridge Univ. Press, 1992).
8. Wigley, T. M. L. & Raper, S. C. B. *Nature* **357**, 293–300 (1992).
9. Mitchell, J. F. B., Manabe, S., Tokioka, T. & Meleshko, V. in *Climate Change 1992. The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. & Ephraums, J. J.) 131–172 (1990).
10. Gates, W. L., Mitchell, J. F. B., Boer, G. J., Cubasch, U. & Meleshko, V. P. in *Climate Change 1992. The Supplementary Report to the IPCC Scientific Assessment* (eds Houghton, J. T., Callander, B. A. & Varney, S. K.) 97–134 (Cambridge Univ. Press, 1992).

Snap judgement

TELEVISION, says Daedalus, has a strange and terrible authority over its victims. It dictates their hopes and dreams, their sense of what life ought to be, even their sense of what life is. For narcotized millions, the characters in fictional soap operas are more real than any human family. Politics is infected too — the whole US Somalian adventure was triggered, not by considerations of policy, but by television images chosen and assembled by news editors. The moving picture, which began so promisingly, is now a major social curse. Proliferating by video, cable and satellite, its disruptive fantasies are even infecting the vulnerable Third World.

Television technology rests, of course, on an illusion: the fusion of 25 still images per second into an apparently moving picture. This image-fusion rate is not immutable. In species such as bees it is about 300–400 Hz. If bees can do it, says Daedalus, so can we. Last week DREADCO's biochemists were trying to extend the persistence of vision. They are now trying to reduce it. By speeding the diffusion of crucial ions through the membranes of the visual pathway, DREADCO's 'Quickview' tablets will raise our image-fusion rate to 1 kHz or higher.

Users of 'Quickview' will find new delight in the visual world. They will resolve previously blurred motions — the whirl of a dancer, the wing beat of a butterfly — into sharp, detailed movements. At work and on the road, their high-speed sight will detect and avert accidents. Best of all, they will be safe from television and film. Television will be a set of lines flashing down the screen; film will be reduced to a jerky sequence of stills. In neither case will there be any illusion of a moving image. Parents will rush to dose their children with 'Quickview' to cure their addiction to video nasties and computer games.

But how to cure the parents themselves? Daedalus will publish fake reports that 'Quickview' counters the effects of food-additive allergies, electromagnetic fields, passive smoking, and similar fashionable perils. Politically correct governments will react by putting 'Quickview' in bread or salt or water, as calcium, iodine and chlorine are now. At a stroke, the moving image will vanish from our lives. Its hapless junkies will be returned to the real world.

One snag is that most computer screens will become meaningless as well. But computer users don't need a moving image; they can work perfectly well with a succession of stills. Re-equipped with long-persistence or non-scanning LCD screens, they will hack valiantly on.

David Jones

Sex hormone response to alcohol

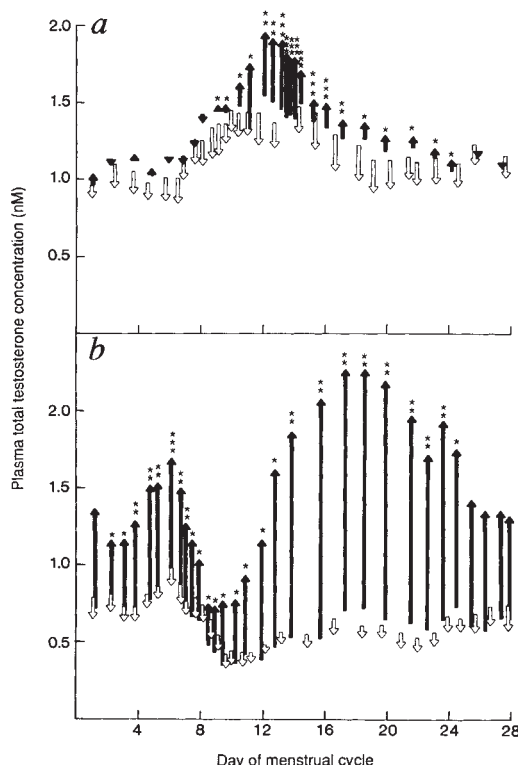
SIR — It is well known that acute alcohol increases subjective feelings of sexual arousal, excitement and desire¹. The controversial suggestion has been made that the sex steroid testosterone could be one factor involved in the basic regulation of these feelings and behaviour^{2,3}. It has been shown, however, that acute alcohol ingestion, if anything, tends to decrease blood testosterone levels in normal healthy men⁴. This makes it unlikely that heightened sexual arousal would be mediated by acute testosterone changes induced by alcohol intake in men, and, so far, testosterone has not been linked to the increased sexual excitement observed in women. But such a linkage could be possible, as we have found that alcohol, even at a very low dose, increases blood plasma total testosterone concentrations very quickly in women but not at all in men.

We gave healthy premenopausal women a low dose (0.34 g kg⁻¹) of alcohol, or placebo juice, on two separate occasions starting at 6 p.m. Testosterone levels decreased slightly during placebo conditions, but alcohol increased the testosterone concentrations (see figure). In the subjects not using contraceptives, the testosterone changes were most pronounced during the ovulatory phase, when basal testosterone was also at its maximal concentration. In the group using oral contraceptives, and thus having substantially lower (average 50%) basal levels than non-users, marked elevations were observed after alcohol intake, with the maximal changes appearing around days 17–20. During these days, the testosterone levels after alcohol exceeded those during corresponding days in subjects not using contraceptives. The steep decrease of basal testosterone levels during days 6–10 was probably due to the start of the daily ingestion of contraceptive pills after menstruation.

In corresponding experiments on 48 young healthy men, no effects on testosterone were observed as a function of alcohol. In another study, the effects of the alcohol dose (0.34–1.02 g kg⁻¹) and time after drinking (40–150 min) were investigated in 10 women using oral contraceptives. Alcohol markedly increased testosterone levels but no dose or time effects were observed.

Acute alcohol-induced elevations of testosterone levels in women have not

been reported previously. This could be due to the fact that the effect is particularly clear only in women taking oral contraceptives and in non-users during the ovulatory phase. Moreover, previous reports have included fewer than 10 subjects per study^{5–7}, and the present work is, to our knowledge, the first major investigation on this topic. Further, detection of the clear effects of alcohol was contingent



Average placebo (open arrows) and alcohol-induced (closed arrows) testosterone changes during different days of the menstrual cycle. *a*, Women not taking the contraceptive pill; *b*, women taking the pill. Each arrow marks the mean ($n=7$) of the average changes during 1 and 2 hours after the start of drinking. The diagram is arranged so that each arrow shares 6 data points with arrows adjacent to it, 5 with arrows 1 position removed, 4 with arrows 2 positions away, and so on, and none with arrows 7 or more positions away from it. Matched *t*-tests were used to calculate the significance (* $P<0.05$, ** $P<0.01$, *** $P<0.001$) between the placebo and alcohol-induced changes. Further experimental details are available from the authors by request.

on subjects receiving a placebo to control for confounding circadian effects, and on the phase of the menstrual cycle and the use of oral contraceptives. Nevertheless, the present results are consistent with one earlier study in which no testosterone changes occurred after alcohol, despite a decrease after the placebo, during days 21–22 of the cycle⁷.

The cause of the alcohol-mediated testosterone elevation in women remains unknown. The fact that no clear dose or

time effects were observed may suggest that alcohol metabolism is involved in testosterone increase.

Although the levels of free testosterone and/or its metabolites are thought to be the active component, we measured total testosterone levels in the present study because the effects in the initial studies were so clear in women. Free testosterone and sex-hormone-binding globulin levels should also be assessed in future, especially in men, because of the failure to find any significant changes in their total testosterone concentrations after they had taken low doses of alcohol.

Our results, demonstrating a testosterone elevation in women but not in men, may provide a physiological explanation for subjective discrepancies^{8–10} in sexual excitement reported by men and women after drinking. While subjective excitement co-varies with penile tumescence in men, both effects being suppressed at higher alcohol concentrations⁹, elevated sexual arousal remained in conjunction with increasing estimates of intoxication in women, although vaginal responsiveness was also reduced^{8,10}. There is as yet no convincing explanation for these gender differences. The alcohol-mediated testosterone elevation discovered here may provide at least a partial answer, assuming that testosterone and/or its active metabolite(s) could trigger feelings of sexual excitement.

C. J. Peter Eriksson

Biomedical Research Center,
Alko Ltd, PO Box 350,
00101 Helsinki, Finland

Tatsushige Fukunaga

Department of Legal Medicine,
Shiga University of Medical Science,
Otsu 520-21, Japan

Ralf Lindman

Department of Psychology,
Åbo Akademi University,
20500 Åbo, Finland

1. Crowe, L. C. & George, W. H. *Psychol. Bull.* **105**, 374–386 (1989).
2. Bancroft, J. J. *Br. med. Bull.* **37**, 153–158 (1981).
3. Alexander, G. M. & Sherwin, B. B. *Psychoneuroendocrinology* **18**, 91–102 (1993).
4. Välimäki, M. & Ylikahri, R. *Alcohol Alcohol.* **18**, 313–320 (1983).
5. McNamee, B. *et al. Br. J. Addict.* **74**, 316–317 (1979).
6. Becker, U. *et al. Drug Alcohol Depend.* **22**, 141 (1988).
7. Välimäki, M., Härkönen, M. & Ylikahri, R. *Alcohol. clin. exp. Res.* **7**, 289–293 (1983).
8. Wilson, G. T. & Lawson, D. M. *J. abnorm. Psychol.* **85**, 489–497 (1976).
9. Wilson, G. T. & Lawson, D. M. *J. abnorm. Psychol.* **85**, 587–594 (1976).
10. Wilson, G. T. & Lawson, D. M. *J. abnorm. Psychol.* **87**, 358–367 (1978).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references.

Heading detection from optic flow

SIR — The question of whether humans can visually detect the direction of egomotion in the presence of a simulated eye rotation has been the subject of recent debate^{1–3}. Psychophysical studies using slightly different stimulation procedures have yielded very different results when eye rotations in excess of 1 degree s^{−1} were simulated. But the flow stimuli used in these studies do reveal a qualitatively different pattern of motion when the flow actually arriving on the retina is analysed (Fig. 1). In situations where humans succeed^{1,3}, the retinal flow field is overall centrifugal in structure, whereas in situations where humans cannot correctly detect their direction of heading², it is not.

It has been found independently in the visual cortex of cats and monkeys that most cells in some areas specialized for motion processing favour movements away from the fovea or area centralis^{4,5}. This centrifugal bias is present in cortical areas PMLS (cat) and MT (monkey), and has been hypothetically linked to the processing of optic flow fields during egomotion.

To see whether such an anisotropic distribution of motion detectors could be related to the differences observed in human psychophysics, we introduced a

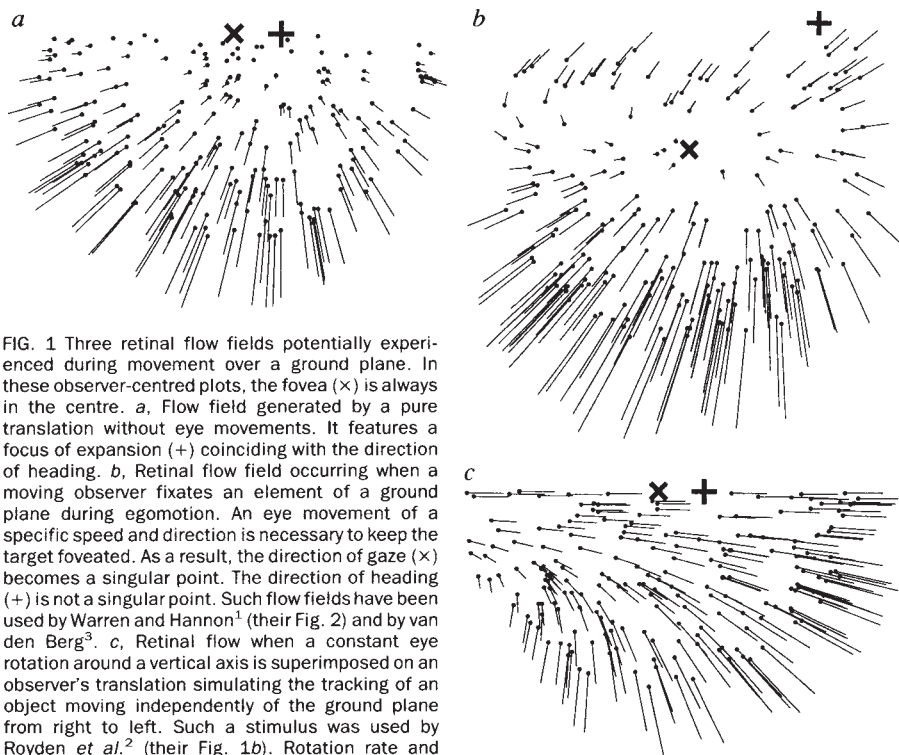
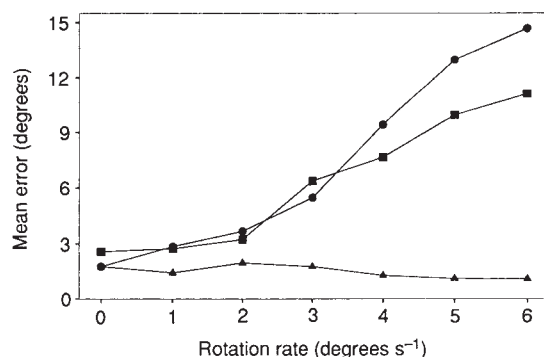


FIG. 1 Three retinal flow fields potentially experienced during movement over a ground plane. In these observer-centred plots, the fovea (x) is always in the centre. *a*, Flow field generated by a pure translation without eye movements. It features a focus of expansion (+) coinciding with the direction of heading. *b*, Retinal flow field occurring when a moving observer fixates an element of a ground plane during egomotion. An eye movement of a specific speed and direction is necessary to keep the target foveated. As a result, the direction of gaze (x) becomes a singular point. The direction of heading (+) is not a singular point. Such flow fields have been used by Warren and Hannon¹ (their Fig. 2) and by van den Berg³. *c*, Retinal flow when a constant eye rotation around a vertical axis is superimposed on an observer's translation simulating the tracking of an object moving independently of the ground plane from right to left. Such a stimulus was used by Royden *et al.*² (their Fig. 1*b*). Rotation rate and translation speed in *c* are identical to *b*, only the direction of the eye movement differs. As in *b*, the direction of heading (+) is not a singular point. The overall structure of the flow field in *c*, however, is quite different from the one in *a* and *b*. In *c*, the individual movement vectors are arranged in a curved fashion, suggesting a movement along a curved path, whereas in *a* and *b* the flow fields have an overall centrifugal structure. Visual heading detection in humans works well with stimuli as in *a* and *b* (ref. 1, 3), but fails with flow fields of type *c* (ref. 2).

FIG. 2 Mean errors yielded by a network model of heading detection for different flow field stimuli and varying amounts of rotation. When the eye rotation is due to the fixation of a ground plane element, as in Fig. 1*b*, the heading error is low and independent of rotation rate (triangles). When eye rotation around a vertical axis is superimposed, as in Fig. 1*c*, the error rises drastically as the rotation rate increases (circles). The same is true when instead of a ground plane the observer moves through a random cloud of dots at a speed of 0.5 m s^{−1} (squares). In this case, too, the retinal flow field is not radial but rather lamellar for higher rotation rates. Visual field diameter was 34°, and observer speed in the first two simulations was 1.9 m s^{−1}. The network model is a biologically plausible implementation of a least-square algorithm for heading detection⁶. In the input stage of the network the retinal flow is encoded by a population of direction-selective motion detectors which are modelled after cells in visual cortical areas PMLS in cat or MT in monkey. The population contains an anisotropy such that the preferred directions of the motion detectors are arranged in a centrifugal or circular fashion around the fovea, and no centripetal preferences are present. The outputs of the motion detectors feed into a second stage in which the direction of heading is recovered. Cells in this layer form populations that test the compatibility of the measured retinal flow field with a specific preferred heading direction. The best matching direction can finally be detected by a simple winner-take-all scheme, or another more elaborate design which operates on the population activities.



similar anisotropy in a proposed network model of heading detection. The model uses motion detectors similar to cells in PMLS and MT to encode the retinal flow field and in a second stage recovers the direction of heading with populations of cells designed to implement a least-square heading detection scheme. In computer simulations, our model cells respond to various combinations of expanding, contracting, rotating or translating random-dot patterns⁶. Interestingly, similar response properties have been described for

a subclass of cells in monkey visual area MSTd, which is a stage of the visual motion pathway in monkey cortex subsequent to area MT^{7–10}.

When we established an anisotropic distribution of direction preferences in the network's input layer by removing all cells with preferred direction towards the fovea, the network exhibited differences in detecting the direction of heading from different flow field stimuli which were similar to those in humans (Fig. 2). This suggests (1) that the observed differences in humans could be related to a specialization of the heading detection system to centrifugal flow fields; and (2) that an anisotropic distribution of motion detectors seems to be a reasonable design for ground-living mammals, for which the fixation of a stationary object somewhere near the movement trajectory is probably the most common condition of egomotion.

Nevertheless, accurate heading detection might also be of vital importance when a moving target is tracked instead of a stationary object. The experiments of Royden *et al.*² indicate that this is a more difficult task, and we might not be able to perform it had we not the benefit of extraretinal information about our eye

- Warren, W. H. Jr & Hannon, D. J. *Nature* **336**, 162–163 (1988).
- Royden, C. S., Banks, M. S. & Crowell, J. A. *Nature* **360**, 583–585 (1992).
- Van den Berg, A. V. *Nature* **365**, 497–498 (1993).
- Rauschecker, J. P., von Grünau, M. W. & Poulin, C. J. *Neurosci.* **7**, 943–958 (1987).
- Albright, T. D. *Vis. Neurosci.* **2**, 177–188 (1989).
- Lappe, M. & Rauschecker, J. P. *Neural Comp.* **5**, 374–391 (1993).
- Duffy, C. J. & Wurtz, R. H. *J. Neurophysiol.* **65**, 1329–1345 (1991).
- Graziano, M. S. A., Andersen, R. A. & Snowden, R. J. *Neurosci.* **14**, 54–67 (1994).
- Orban, G. A., Lagae, L., Verri, A., Raiguel, S., Xia, D., Maes, H. & Torre, V. *Proc. natn. Acad. Sci.* **89**, 2595–2599 (1992).
- Tanaka, K. & Saito, H.-A. *J. Neurophysiol.* **62**, 626–641 (1989).
- Gibson, J. J. *The Senses Considered as Perceptual Systems* (Houghton Mifflin, Boston, 1966).

movements. Undoubtedly, extraretinal information is used by the human heading detection system to make otherwise unsolvable tasks less ambiguous, as has been previously demonstrated for movements towards a vertical plane¹. We do not intend to suggest that humans (or animals in general) do not use extraretinal information; our main interest is rather the possible adaptation of the visual system to frequently experienced natural flow fields in the sense of Gibson's 'visual ecology'¹¹.

Markus Lappe

Departments of Zoology

and Neurobiology,

Ruhr University Bochum,

Bochum,

D-44780 Germany

Josef P. Rauschecker

Laboratory of Neurophysiology,

NIMH,

Poolesville,

Maryland 20837, USA

Lunar crater chains

SIR—The recognition¹ that nearly a dozen crater chains on the jovian moon Callisto may have been created by comets disrupted during passage inside Jupiter's Roche limit has prompted us to search for crater chains on our Moon that may have been created by comets similarly disrupted near the Earth. A careful search of the available global photography² of the Moon yielded two chains (see table) that are not obviously secondaries from any large crater or basin, nor are endogenic. Both chains are on the lunar nearside, as expected for comets split during a close pass by the Earth, and range in length from about 50 km for the well-known

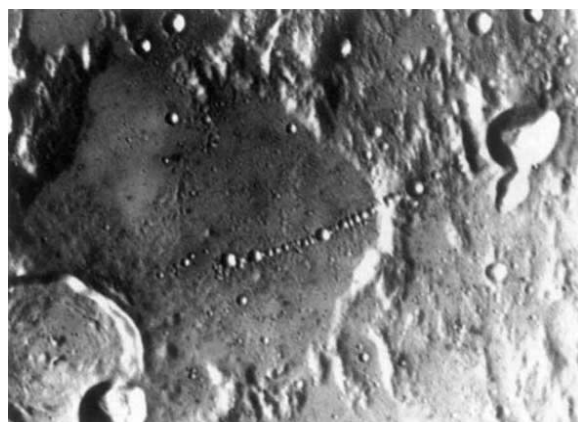
elongated in the direction of the chain.

Although the Davy chain is a relatively fresh, post-Imbrian feature⁵, the Abulfeda chain is more degraded and is assigned an Imbrian age⁵. The disparity in the lengths of these chains might at first suggest that they do not have similar origins. However, a simple model for tidal splitting¹ that assumes a radial breakup at perigee shows that the length of the chain is a linear function of the parent size. Because the craters in the Abulfeda chain are roughly five times larger than those in the Davy chain, it is not unreasonable to suggest that the Abulfeda chain is roughly five times longer.

The tidal stresses experienced by an object approaching within the Earth's Roche limit are actually larger than those approaching Jupiter at a similar relative distance. Using a simple estimate of the tidal stress⁶, σ_{tide} is given in terms of the comet density ρ , mass of the Earth M_E , radius of the comet r_c and distance of approach R .

$$\sigma_{\text{tide}} \simeq 0.15 \frac{\rho G M_E}{R^3} r_c^2 = 0.63 \frac{G \rho \bar{\rho}}{(R/R_E)^3} r_c^2$$

The second equation is written in terms of the planet's mean density $\bar{\rho}$, and it is clear that for approach at a given fraction of the radius of any planet, the tidal force is proportional to $\bar{\rho}$. Because the Earth's mean density is about five times that of Jupiter, the tidal forces on an object approaching to, say, 1.5 radii are correspondingly larger. The tidal stress on an Earth-approaching comet is thus about $\sigma_{\text{tide}} = 6.8 \times 10^{-4} r_c^2$ (km), where the stress is in bar. Despite these relatively larger stresses, the inferred strength of comets is still very small (measured



The remarkably regular shape, spacing and alignment of the craters in the 47-km-long Davy crater chain is unmatched anywhere else on the lunar surface. The unimpeded continuation of the craters across the floor and up the east wall of crater Davy Y is inconsistent with an endogenic origin. The linear dunes separating many of the craters indicate nearly simultaneous impacts⁴. Portion of Apollo 16 metric frame 2198.

the Moon's disk, the observation of 1 ± 1 crater chains on the moon implies about $2 \times 10^5 \pm 2 \times 10^5$ split comets overall, assuming that comets approach the Earth in random directions. Over 4 Gyr, this suggests a recurrence interval of at least 2×10^4 years, assuming a constant flux in post-Imbrian time. If 2×10^5 comets passed between the Earth's surface and its Roche limit, then about 0.19 of these comets should have struck the Earth itself (neglecting velocity-dependent gravitational focusing). Over 4 Gyr this implies a cratering rate of about 10^{-5} cometary impacts per year, which is much larger than the estimate in ref. 7 of 10^{-7} per year of comets with diameters greater than 2.5 km. But the projectiles that made the craters in the Davy chain were probably no more than 100 m in diameter, using current scaling theory⁸. In this case the parent would only have had a diameter of about 300 m, and if the cumulative number of comets equal to or greater than diameter D scales as D^{-2} , as suggested by cometary brightness distributions⁹, then the estimated flux of 300 m diameter comets is at least consistent with the estimate of Shoemaker *et al.*⁷. Another

LUNAR CRATER CHAINS THAT MAY HAVE BEEN MADE BY TIDALLY SPLIT COMETS

Name	Latitude	Longitude	Length (km)	Crater diameter range (km)	Number of craters	Lunar orbiter frame
Davy	-10°	-9°	47	1-3	~ 23	IV-108-H2
Abulfeda	-15°	15°	200-260	5-13	~ 24	IV-89-H2

Davy chain³ (see figure) to 200–250 km for a chain near the crater Abulfeda.

Secondary crater chains on the Moon are common, but such chains are usually radial or nearly radial to a large crater or basin, occur in the near vicinity of other such chains, and have raised rims with a characteristic 'chevron' imprint between craters that points back to the primary³. Individual secondary craters are frequently irregular in shape. Much rarer are chains, such as the northwest arm of Rima Hyginus or a small chain near the crater Beer, that are obviously of endogenic origin. These craters exactly fill the interiors of rilles or grabens and are often

in millibars for comets of a few kilometres radius) and suggests a model for comets in which a few mechanically strong subunits are held together by self-gravity alone.

If the two crater chains near Davy and Abulfeda were created by tidally split comets, a rough bound on the cometary flux may be deduced. As the Abulfeda chain was probably created during the high heavy bombardment impact flux before 3.8 Gyr ago, we will estimate the current flux only from the post-Imbrian Davy chain. Of course, the statistics of 1 ± 1 crater are very poor, but one chain is all we have. From the ratio of the areas of a sphere enclosing the Moon's orbit and

- Melosh, H. J. & Schenk, P. *Nature* **365**, 731–733 (1993).
- Bowker, D. E. & Hughes, J. K. *Lunar Orbiter Photographic Atlas of the Moon* (Spec. Publ. 206, NASA, Washington DC, 1971).
- Masursky, H., Colton, G. W. & El-Baz, F. *Apollo Over the Moon: a View from Orbit* (Spec. Publ. 326, NASA, Washington DC, 1978).
- Oberbeck, V. R. & Morrison, R. H. *The Moon* **9**, 415–455 (1974).
- Wilhelms, D. E. & McCauley, J. F. *Geologic Map of the Near Side of the Moon* (US Geol. Surv. Washington DC, 1971).
- Scotti, J. & Melosh, H. J. *Nature* **365**, 733–735 (1983).
- Shoemaker, E. M., Wolfe, R. F. & Shoemaker, C. S. in *Global Catastrophes in Earth History* (eds Sharpton, V. L. & Ward, P. D.) 155–170 (Geol. Soc. Press, Boulder, Co., 1990).
- Melosh, H. J. *Impact Cratering: A Geologic Process* (Oxford Univ. Press, New York, 1989).
- Delsemma, A. H. in *Symposium on the Diversity and Similarity of Comets 19–20* (Eur. Space Agency, Brussels, 1987).

possibility, which we cannot adequately evaluate, given the small number of crater chains, is that the Davy and perhaps the Abulfeda chains were created by tidally disrupted 'rubble pile' asteroids.

H. J. Melosh

E. A. Whitaker

Lunar and Planetary Laboratory,
University of Arizona,
Tucson, Arizona 85721, USA

Statue enigma

SIR — Ragge and Munier in Scientific Correspondence¹ presented a Hellenistic statue displaying multiple skin nodules. While I agree with their diagnosis of neurofibromatosis type 1, I think that it is more likely to have been a votive offering than a teaching aid.

Ancient votive offerings did include representations of pathological states, involving both the extremities (the most famous being the oversized leg with a varicose vein from Athens²) and the torso (for example a female statue with ascites from Cos³). Even the Etruscan terracottas showing normal and pathological visceral anatomy in a fairly detailed manner clearly were votive offerings⁴ and did not form part of an anatomical museum. Therefore, there is a lot of evidence for the traditional interpretation of the neurofibromatosis statue as a votive offering, but little to support the view that it was a teaching aid.

The hypothesis that a certain group of the "grotesque" terracotta figures, which were found at Smyrna and various other centres of the Hellenistic world and to which this statue bears resemblance, served as teaching aids for the local medical schools^{5,6} is itself questionable. The existence of pathological collections for teaching purposes would be alien to all that we know about ancient medical education. It would be surprising that Galen, who attended the medical school situated at Smyrna⁷, should not mention a collection which would without doubt have been of enormous interest to him.

With regard to the style of the statues, the element of caricature often prevails over anatomical exactitude⁸, which would be rather odd for teaching material. The only figures for which the 'teaching-aid' hypothesis may be considered are those showing acute problems amenable to surgical intervention, such as the scrotal enlargement or the acute dyspnoea due to a foreign body in the airways (Figs D 1203, D 1211-2 of ref. 8).

Chronic conditions of the skin were regarded as incurable by the methods of scientific medicine and therefore a domain of both popular and temple medicine⁷. How could the man who suffered from neurofibromatosis make it clearer to the god Asclepius what he wanted to be

relieved from, than by a sculptural representation of the remarkable features of his disease?

Unfortunately, the statue has apparently disappeared from the Meyer-Steineg Collection at Jena University⁹. The complete survey of this collection, announced in ref. 9, might bring further evidence. But in my view, the statue was not an "early three-dimensional teaching aid" (ref. 1), but a propitiatory offering, given by a man who hoped to be cured by the god from his deforming condition.

David E. J. Linden

Christ Church,
Oxford OX1 1DP, UK

1. Ragge, N. K. & Munier, F. L. *Nature* **368**, 815 (1994).
2. Margotta, R. & Lewis, P. *An Illustrated History of Medicine* p. 55 (Hamlyn, Feltham, 1968).
3. Meyer-Steineg, T. *Darstellungen normaler und krankhaft veränderter Körperteile an antiken Weihgaben*, *Jenaer medizin-historische Beiträge* **2**, 18 (Verlag von Gustav Fischer, Jena, 1912).
4. Decoufle, P. *La notion d'ex-voto anatomique chez les Etrusco-Romains* 15–32 (Collection Latomus, vol. 72, Bruxelles, 1964).
5. Laumonier, A. *Bulletin de correspondance hellénique* **70**, 318 (1946).
6. Gourevitch, D. & Gourevitch, M. *Presse méd.* **55**, 2752 (1963).
7. Ilberg, J. in *Antike Medizin* (ed. Flashar, H.) 366–367 (Wissenschaftliche Buchgesellschaft, Darmstadt, 1971).
8. Besques, S. *Catalogue raisonné des figurines et reliefs en terre cuite grecs, étrusques et romains. Époques hellénistique et romaine* 3 (Editions des Musées Nationaux, Paris, 1971/2).
9. Künzl, E. in *Le Latin médical* (ed. Sabbah, G.) 112 (Publications de l'Université de Saint-Etienne, 1991).

As planets go by

SIR — The search for protoplanetary disks around normal main-sequence stars like our Sun has gained wide interest because it offers the unique opportunity to study one of the preconditions for extra-terrestrial life, namely the formation of planets. The discovery of a far-infrared excess around Vega, β Pictoris and a few other nearby stars¹ was interpreted as having originated from circumstellar dust grains with sizes of $\sim 50 \mu\text{m}$, which is considerably larger than the constituents of the interstellar medium. This fits into a picture where circumstellar grains grow owing to collisions during their orbit. Because the wavelength of maximum emission of a dust grain increases with its size, observations at millimetre wavelengths can provide evidence for still larger particles. Therefore, we observed² some of these stars at $\lambda = 1.3 \text{ mm}$ in an 11-arcsec beam and derived grain-size distributions where particle diameters up to 1 cm were required to explain the data. Measurements in a 24-arcsec beam indicated that the 1.3-mm emission is extended on a scale of several 100 AU³.

Stern *et al.*^{4,5} have reported the detection of a 200-AU extended disk-like structure by mapping the area around Fomalhaut (α PsA) at 1.3 mm. Such a discovery would have been the second case (after β Pictoris) of a disk and the first

direct evidence for a spatially resolved, flattened structure at 1.3 mm around a main-sequence star. The 1.3-mm contour map obtained by these authors, however, is noisy and also shows emission from outside the postulated disk area⁵. Similarly, the inferred 1.3-mm flux densities from the central 11 arcsec at the stellar position of 50 mJy (ref. 4) and $32 \pm 12 \text{ mJy}$ (ref. 5) conflict with our result² of $7.3 \pm 2.2 \text{ mJy}$ and with that of Mannings and Emerson⁶, who found a 3σ upper limit of 24 mJy. For that reason we decided to clarify the situation by repeating the 1.3-mm observations on Fomalhaut.

Between 6 and 9 April 1994, we used the MPIfR seven-channel bolometer array⁷ at the IRAM 30-m telescope and performed on-off measurements at the position of the star. The beam separation, provided by the wobbling secondary mirror was 32 arcsec. The weather was stable, with a zenith opacity between 0.08 and 0.13 at 1.3 mm; Uranus served as a calibrator. In this way we obtained a central (11 arcsec) flux density of $6.7 \pm 2.1 \text{ mJy}$, in perfect agreement with our previous result². For the total flux in the six surrounding beams of the hexagonal array configuration we derive a 3σ upper limit of 26.4 mJy which is also consistent with our measurement of $21.0 \pm 2.5 \text{ mJy}$ within a 24-arcsec area³; these data, however, are not compatible with the results of Stern *et al.*⁵.

To investigate the extended emission claimed by Stern *et al.*⁵, we performed three maps at the position of Fomalhaut, each $180 \times 96 \text{ arcsec}$ in size. The figure shows our final co-added map which is drawn for direct comparison with Stern *et al.*⁵ over an identical field and with the same contour levels. Our map has a factor of two better sensitivity ($1\sigma \text{ r.m.s.} = 6.3 \text{ mJy}$) and shows no hint of extended emission. We therefore conclude that the disk claimed by Stern *et al.*⁵ must be erroneous owing to observational artefacts. There is no significant dust emission from Fomalhaut above the limits given by us^{2,3}.

Rolf Chini

Endrik Krügel

Ernst Kreysa

Robert Zylka

Max-Planck-Institut für
Radioastronomie,
53121 Bonn, Germany

1. Aumann, H. H. *et al.* *Astrophys. J.* **278**, L23–L27 (1984).
2. Chini, R., Krügel, E. & Kreysa, E. *Astr. Astrophys.* **227**, L5–L11 (1990).
3. Chini, R., Krügel, E., Shustov, B., Tutukov, A. & Kreysa, E. *Astr. Astrophys.* **252**, 220–228 (1991).
4. Stern, S. A., Festou, M. C. & Weintraub, D. A. *IAU Circ. No.* 5732 (1993).
5. Stern, S. A., Festou, M. C. & Weintraub, D. A. *Nature* **368**, 312–314 (1994).
6. Mannings, V. & Emerson, J. P. *IAU Circ. No.* 5786 (1993).
7. Kreysa, E., Lemke, R., Haslam, C. G. T. & Sievers, A. *Astr. Astrophys.* (in the press).

■ See the correction by Stern *et al.* page 766.

Nowhere to hide

Dorothy Nelkin

Who Owns Information? From Privacy to Public Access. By Anne Wells Branscomb. BasicBooks: 1994. Pp. 241. \$25.

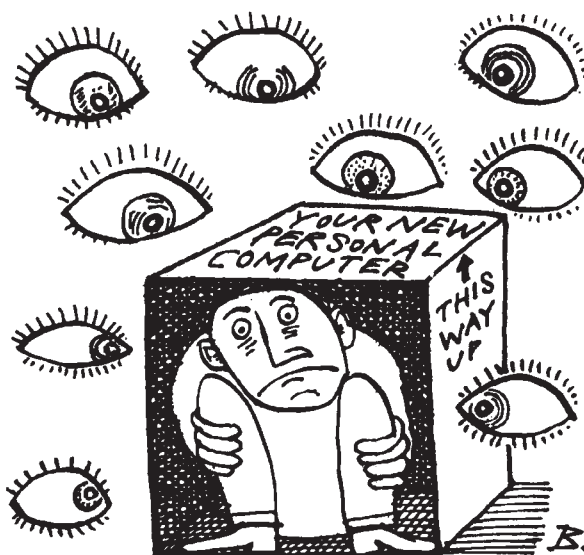
DESPITE its profound social impact, the information revolution has by and large been exempt from the public disputes that so frequently surround technological change in the United States. The technologies that make it possible to manipulate, store and gain access to information have intruded on privacy, threatened civil liberties and imposed on many rights; yet they have evoked surprisingly few expressions of public concern. In *Who Owns Information?*, a useful, accessible and often entertaining book, Anne Branscomb, an attorney and legal scholar, has laid out the social and legal dilemmas posed by communication and information technologies, suggesting the inadequacy of present laws and regulations to deal with their complex implications. She is especially concerned with the conflict between individual privacy and public access to information — the blurring of the lines between public and private domains of information.

Branscomb poses nine questions. Who owns your name and address? Your telephone number? Your medical history? Your image? Your electronic messages? Video entertainment? Religious information? Computer software? And government information? In each case, she provides examples to illustrate how information systems such as direct mail and telemarketing work. She describes the commercial and political value of personal information, providing titbits of amazing and discouraging information for those in the United States. When we file a post-office change-of-address card, it becomes available to direct-mail marketers. When we dial an 800 number to ask for information, we are likely to be placed on a telemarketing list. When we subscribe to a magazine, use a credit card, send an e-mail message or visit a doctor, records become available that may be accessed by organizations with commercial interests.

Branscomb shows how technological advances put a new spin on familiar dilemmas. In the academic disputes over the ownership of the Dead Sea Scrolls, biblical scholars working with computers and photographic images are carrying on the traditional struggles among priests over custody of religious dogma. Debates about computer software add new di-

mensions to intellectual property claims. The legal status of our information assets, she argues, is alarmingly vague.

Three areas of law bear on the ownership of information in the United States: First Amendment rights to protect freedom of the press, intellectual property rights to encourage innovation and privacy rights to protect individuals. These were established well before recent technological changes. Branscomb therefore



argues for "a coherent effort to rationalize the legal infrastructure (*sic*) of the information age" on the basis of a clear set of ethical values and guiding principles that Americans would be prepared to follow.

But what values and principles would, in fact, be acceptable to most Americans? Despite the intrusiveness of information technologies, criticisms have mainly come from an élite — ethicists, sociologists, lawyers and other professionals. There is a striking absence of organized public concern. Indeed, Americans seem to care little about privacy. Although computerized databanks allow bureaucratic authorities to have easy access to personal information about credit ratings, school performance, housing, medical history, tax status and even genetic profile, few people seem outraged. More and more information is available to employers, insurers, advertizers, banks, school systems and other institutions that exercise control over our lives, yet we accept new information technologies as useful and benign. There is little public outcry against the surveillance gimmicks that photograph us in department stores, or

the 'muzak' that intrudes on our privacy in supermarkets and airports. There is no apparent public concern about the privacy implications of the Clinton administration's proposal for a universal health card, even though it would contain, in accessible form, the complete health history of every American.

Popular magazines and television and radio talk shows are full of lurid and embarrassing personal confessions, and the audiences seem to relish intrusions on privacy. To an amazing degree, people talk about their own personal problems in public. Indeed, relinquishing privacy is seen as a way to solve personal problems; far from demanding privacy, Americans let it 'all hang out'. Perhaps this explains why, despite their obvious intrusions, information technologies have not been resisted.

The rapid development of computerized databanks, allowing a striking level of institutional control over individuals, has evoked some professional concern. Yet it has not brought about a great deal of public resistance. People see information technologies through their computers and fax machines, which have been marketed as a way of empowering and liberating the individual — of expanding individual choice. Perhaps no industry has been more successful than the information-technology industry at turning the latest gimmick — the extra megabyte, the latest fax machine, and now the videophone — into a dire necessity. The gadgets seem to give people more, not less, control, blinding them to potential abuses.

Nor do we seem to care that along with the information highway comes the risk of hegemonic control over the messages we receive from the media. We welcome the advances in information technology that have brought innumerable television channels as 'pluralism', ignoring their role in the corporate manipulation of consumer taste and cash flow. And we welcome proposals for electronic democracy, failing to see the difference between the inundation of information and reflective political exchange.

Branscomb nicely sorts out the critical legal and ethical dilemmas generated by new information technologies and makes an effective case for better laws. But she underestimates the political will to control technologies that are widely viewed more as symbols of progress and icons of US ingenuity than as commercial and political products that threaten individual rights. □

Dorothy Nelkin is in the Department of Sociology, New York University, 269 Mercer Street, New York, New York 10003, USA.



AMATEUR astronomers not only hunt for comets (see below); they also flock to sites around the world to observe eclipses. This photograph of an annular eclipse is from *Chasing the Shadow* by Joel Harris and Richard Talcott, an observer's guide to solar eclipses through to the year 2020. Kalmbach, \$18.95 (pp. 160, pbk).

Debris from space

Don Yeomans

The Quest for Comets: An Explosive Trail of Beauty and Danger. By David H. Levy. Plenum: 1994. Pp. 280. \$23.95.

SPENDING long lonely nights searching for new comets is a surprisingly competitive pursuit and only the most persistent succeed. Experienced observers spend an average of 200 search hours per discovery. But the rewards are great. Not only will a newly discovered comet be available for scrutiny by other astronomers, but the name of the discoverer is attached to the comet from then on. There is a certain immortality in having your name associated with an object that will normally outlive you by several million years.

David Levy is one of the more successful comet hunters, and his engaging writing style and love for his science is contagious. Using historical and personal anecdotes, he explains why he devotes much of his life to pursuing the subtle light of the elusive comets. Because comets are thought to be the least-altered debris remaining from the formation of the outer Solar System four-and-a-half billion years ago, these bits of ice and dust are eagerly studied by planetary scientists. The thread that runs throughout his story concerns the violent collisions of comets with our Earth and Moon as well as with other planets and their moons. Collisions with the early Earth may have spread a veneer of carbon-based molecules and water on the Earth's surface, providing the mixture that allowed life to form. Subsequent cometary collisions with the Earth prob-

ably destroyed large numbers of species and allowed the more hardy species (such as us mammals) to develop further.

The author is at his best when telling the many historical and personal anecdotes that fill this small volume. Comets have been discovered by mistake and have frequently been mistakenly discovered. One comet hunter gave a precious cometary discovery to a friend and another, we learn, found his wife in the classified ads. Much of the book consists of anecdotal accounts of the lives of Levy's current partners in cometary pursuit, Gene and Carolyn Shoemaker. Gene's recollections of the first attempts to explore the Moon during the Explorer, Surveyor and Apollo spacecraft programmes are memorable, as is the evolution of Carolyn from a homemaker to a world-class astronomer. Her cometary findings will soon surpass the record of 37 discoveries made by the early nineteenth-century French astronomer Jean Louis Pons.

One of Carolyn's more recent discoveries, the bizarre comet Shoemaker-Levy 9, was made in March 1994. This comet has split into 21 separate fragments and, rather than orbiting the Sun, they are all in temporary orbits about Jupiter. If we could count each fragment as a separate comet, then Pons's record has already fallen. On the other hand, the suicidal fragments will soon disappear forever when they hit Jupiter during the week of 16–22 July 1994. Using this comet and others as examples, the author has provided us with a personal, well-written narrative on his quest. □

Don Yeomans is in the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109. USA.

Growth of biological thought

Jane Maienschein

A History of the Life Sciences. Second Edition. By Lois N. Magnier. Dekker: 1994. Pp. 496. \$59.75.

It is difficult to tell for whom this book is intended. Lois Magnier sees the volume as useful for a one-term introductory course as well as for general readers and teachers. Yet the price precludes its use as a course textbook, and the choice of publisher makes it fairly inaccessible to a general audience. Perhaps it is best assessed against Magnier's stated objectives: to provide a panoramic view of the forest of concepts and methodologies of the life sciences, while also pausing to examine particular trees in an effort to stimulate interest from historians and biologists.

Magnier writes in a lively style that makes the individual examples accessible and amusing. We learn, for example, that the Enlightenment physician and philosopher La Mettrie attended "a feast given in his honour by a grateful patient. Immediately after eating an enormous quantity of a truffle pastry, La Mettrie fell ill and died. The cause of death was unclear, but Voltaire pronounced it a great occasion since, for once, the patient had killed the doctor". Or of T. H. Huxley: "Physicians told Huxley that his ailing bride was unlikely to live another 6 months; this estimate proved to be wrong by about 50 years". Such stories bring the individuals alive and help to make history more fun. The facts also accord with the best available scholarship, providing reasonably 'accurate' accounts. The trees are presented in focus and with attention to detail.

Yet details, as Magnier realizes, do not alone make history. In trying to present the larger themes, she is less successful. The first four chapters cover familiar ground in the history of science but include too much detail. No coherent story emerges, so that the newcomer is left feeling that there was a lot going on but with no clear understanding of what it all means. It seems unlikely that this section will inspire biologists, and there are already excellent histories of science presenting the same ideas and individuals to historians or more general readers. Magnier has read the literature and provides good reading lists, but the lack of footnotes or annotations means that readers will find it difficult to follow up particular points. But for the initiated the chapters provide a useful reference survey.

The remaining six chapters are devoted to particular subjects or disciplinary areas (generation, physiology, microbiology,

evolution, genetics and molecular biology) and, until the final chapter, dwell on pre-twentieth-century work. Although some of the sections on more recent work may interest scientists, Magner fails to show how ideas, practices and individual approaches to research give way to each other. There is little sense of the forces of change or of how the pieces and players interrelate.

Magner believes that the recent attempt at the social construction of science is unlikely to inspire scientists. So she leans toward an "old-fashioned" historical approach, which, she feels, does a better job of including the science while encouraging young readers to pursue a life in science. Unfortunately, her attempt does not really succeed.

I agree that social construction and rhetorical analysis will not get us far without a clear sense of scientific ideas, methodologies and approaches (recently termed 'practice' by the trend-conscious); but, despite her preferences, Magner does not include enough science. For example, the nine pages on Darwin tell us a lot about his health, family and friends but not much about the idea of evolution by natural selection or how Darwin laboriously put together the vast array of evidence in support of his theory.

The lack of any conclusion reinforces the sense of vignettes all in a row. The final pages point to the Human Genome Project but end abruptly with: "it might be useful to establish a new version of the central dogma, which will simply remind us that 'Knowledge is not equivalent to Wisdom'". Precisely my point about Magner's volume: it provides many useful and enjoyable bits of what can reasonably count as 'knowledge' but it does not move us far toward 'wisdom'. □

Jane Maienschein is in the Departments of Philosophy and Zoology, Arizona State University, Tempe, Arizona 85287-2004, USA.

New in neuroscience

Three books on the brain have recently been published (see also the review of *Images of Mind* on this page):

Localization and Neuroimaging in Neuropsychology edited by Andrew Kertesz, a "comprehensive and thoroughly current review of theory and methodology". Aimed at a specialist audience. Academic Press, pp. 662, \$89.95.

Neural Activity and the Growth of the Brain by Dale Purves, which arises from a lecture series (1992 Lezioni Lincee) intended for a broad audience. Cambridge University Press, pp. 108, £24.95, \$37.95 (hbk), £11.95, \$15.95 (pbk).

Brain Activation by Per E. Roland, a "comprehensive analysis of the functional relationship between brain organization and human behaviour". Particularly suitable for neuroscientists and cognitive scientists. Wiley, pp. 589, \$84.95.

The measure of the mind

Patricia S. Goldman-Rakic

Images of Mind. By Michael I. Posner and Marcus E. Raichle. *Scientific American Library/W. H. Freeman*: 1994. Pp. 257. \$32.95, £19.95.

MUCH has been written about the revolution in molecular neurobiology and the remarkable noninvasive methods of imaging the human brain. But the coming together of neuroscience, cognitive psychology and computerized brain-imaging technology has also led to a revolution in techniques for measuring cognitive processes in humans and for relating them to specific neural systems. Two leading figures in these fields — Michael Posner, a cognitive psychologist, and Marcus Raichle, a neurologist and neuroscientist — have now collaborated to write a clear and beautifully illustrated review of this enticing and exciting multidisciplinary enterprise.

The bells and whistles of imaging methods are perhaps more widely appreciated than the power, simplicity and elegance of the behavioural tools that have enabled psychologists to unravel mental processes. The authors provide valuable examples of the use of simple reaction-time measurements (introduced already in the nineteenth century by F. C. Donders) to derive information about the complexity of the human mind. Other more modern analytical tools and brain-imaging techniques are also lucidly presented. As the authors point out: "Just as the trace of an invisible particle can be seen by the physicists by its effects on other particles, a mental operation becomes visible when it influences the time taken to respond to a specific probe question presented by the experimenter". They make it clear that the capacity to break down cognitive operations into algorithms and subroutines is an essential step toward localizing the structures that carry out the functions. "Without having analyzed beforehand what the components of higher cognitive processes are, there would be little hope of finding where any are located". The resolution of behavioural measurements may possibly supersede the resolution of brain images and associated neuronal structures.

The book is not a reference source and should not be regarded as a critical treatment of cognitive neuroscience. It is not meant for graduate students or research-

ers in the field. Although the selection of topics is generally illuminating and commendable, it reflects the authors' personal choices and experiences. More could have been made of the many anatomical and physiological findings from animal experimentation that form the basis of our knowledge about the neuroscience of sensory, motor and cognitive performance; the subject of memory is avoided (possibly because the area is in a state of flux); some topics (such as mental disorders) are viewed from an oversimplified neurobiological perspective; and the description of cortical layer-



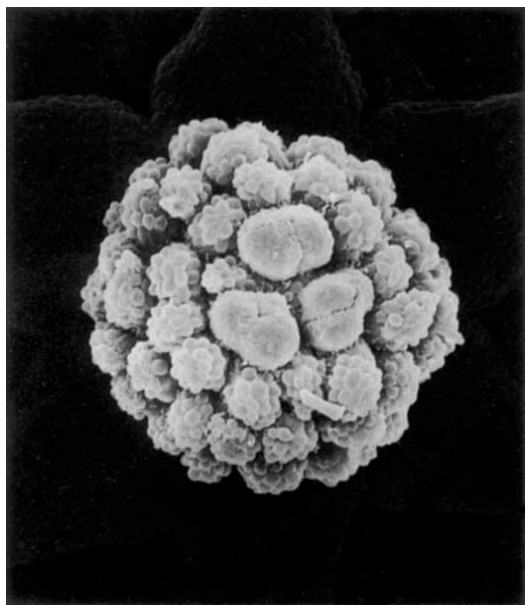
From *Images of Mind*

Self-portraits made by the German artist Anton Räderscheidt in the 9 months following a stroke that damaged his right parietal lobe.

ing in development is mistaken, as shown in the figure on page 183.

For centuries humans have been interested in how their brain carries out its extraordinary intellectual and creative feats, but only now can we directly view our own at work. This book is a celebration of the foundation of cognitive neuroscience, its premises, tools and promise. For the general public, it is more than a coffee-table book; it conveys the fascination and promise of an emerging field. It will be a valuable introduction to cognitive neuroscience for all biologists, neurologists and psychiatrists, as well as undergraduates and anyone interested in knowing how it has become possible to approach the study of mind, not on a philosophical level, but as an experimental science. □

Patricia S. Goldman-Rakic is in the Section of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA.



SEX change — the tiny threadlike plant *Lacandonia schismatica* is unique in that the orientation of its sex organs is reversed: its stamens arise within several rings of pistils. Awarded a family to itself, it was discovered by researchers cataloguing all the vascular plants (some 18,000 species) of Mesoamerica — the southern states of Mexico and the republics of Central America. Two new genera and 104 new species were also found. The project was organized by the Natural History Museum in London, the Universidad Nacional Autónoma de México and the Missouri Botanical Garden. The first volume of *Flora Mesoamericana* (in Spanish) is now published in Mexico.

Steps up the botanical spiral

Martin Ingrouille

Flora Europaea, Volume 1: Psilotaceae to Platanaceae. Second Edition. Edited by T. G. Tutin et al. Cambridge University Press: 1993. Pp. 581. £100, \$200.

The Families of Flowering Plants: Interactive Identification and Information Retrieval. By L. Watson and M. J. Dallwitz. CSIRO: 1994. \$180 (CD-ROM and manual).

Phyllotaxis. By Roger V. Jean. Cambridge University Press: 1994. Pp. 386. £45, \$74.95.

Life Processes of Plants. By Arthur W. Galston. Scientific American Library/W. H. Freeman. Pp. 246. £19.95, \$32.95.

THE production of the second edition of the first volume of *Flora Europaea* is a landmark. The book was first published 29 years ago, and proved a valuable spur to further research. Inevitably the quality of the individual systematic accounts were variable, but the new edition is a great step forward. Many of the accounts have been revised to a high standard by one author, John Akeroyd, and 350 new taxa have been included, hundreds new to science (20 taxa have been deleted).

This might give the impression that everything is cut and dried — and why not, especially for a region with the longest and most extensive history of taxonomy? But consider the last family record in the first volume, the Platanaceae. Not a complicated taxonomic situation. It has a single ancient genus dating back at least to the middle Cretaceous with only two European species: *Platanus orientalis*, the oriental plane, which grows in woodlands

hidden in the steep rocky gorges of the Balkans; and the London plane, *Platanus acerifolia*. *Flora Europaea* still records that the "origin and taxonomic status of this plant have been much discussed but are still uncertain. Some authors consider it to be a hybrid. Others regard it as a cultivar of *P. orientalis*." It used to add in the first edition: "As its origin is unknown there seems at present to be no means of deciding between these possibilities". Well now there is. The whole field of molecular systematics has burgeoned since the first edition and yet this information has scarcely touched the delimitation of taxa recorded in the second edition of the *Flora*. There is work to be done for the European plant taxonomist yet, even in Europe.

The publication of *The Families of Flowering Plants* indicates a new way forward: descriptions, keys and pictures for 563 families on a CD-ROM, which can be accessed 'interactively'. The package requires an MS-DOS-based computer with 640 kilobytes of RAM and 1.5 megabytes of hard-disk space. Installation is easy and the interactive key, which is at the heart of the package, is simple to master with the use of on-screen menus. Subsets of characters, taxa or geographical regions can be chosen. The reader or operator can choose to present the characters in any order. In identifying a specimen, the command 'BEST' gives a list of characters ordered by greatest predictivity. In scoring characters in turn, the number of possible taxa declines, until only a few taxa remain to be compared, or only the correct one. A full description or illustration can then be called up to check

the identification.

The descriptions are recorded in the DELTA format, which maximizes accessibility and the potential for modification. This is the package's greatest strength. It will grow in resolution and sophistication, so its potential for information retrieval and teaching is great. The only danger is that the package is so easy to use that it may lull the unwary into incorrect identifications; in practice, many characters are not readily resolvable, and the primary data have been culled from several sources that no doubt include many incorrect interpretations.

But it is not too grandiose to think that *The Families of Flowering Plants* represents a stage in plant taxonomy as important as the publication of Linnaeus's *Genera Plantarum* in 1737. Like that work, this CD-ROM should provide a stairway for the aspiring botanist, even though it too is imperfect and incomplete; but it is the big first step to a global flora available on disk. Even though the interface is a little awkward, the work is already great fun. A Windows version with greater graphical possibilities is promised.

Phyllotaxis is a step up the botanical spiral staircase. The author presents a united interpretation of the most fundamental patterns of plant morphology, the genetic spiral arrangement of leaves, buds or flowers. This is a marvellous little book. Much of the mathematics is related to Fibonacci numbers and especially the golden ratio, τ ($(\sqrt{5}+1)/2$). "Not only does τ symbolise ratio and proportion, but it also represents one of the oldest symbols of life, τ , the symbol known to ancient Egyptians (the ankh) and ancient Hindus. This book is about one of the most fascinating expressions of life, harmony, and fivefold symmetry." It provides a framework for understanding patterns: what are permitted or more likely, and what are the constraints on developmental or evolutionary change. There are many fascinating examples, side-alleys and speculations. Did you know, for example, that there is a tendency for the genetically determined spiral in palm trees to wind left-handed in the Northern Hemisphere and right-handed in the Southern Hemisphere?

For those wanting more basic facts, *Life Processes in Plants* provides a clear summary of plant physiology, presented with many colour diagrams and the minimum of obscure terminology. One of the latest in the Scientific American Library series, the book is designed for the general reader, although it is by no means superficial. □

Martin Ingrouille is in the Department of Biology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK.

A circumsolar ring of asteroidal dust in resonant lock with the Earth

Stanley F. Dermott, Sumita Jayaraman, Y. L. Xu, B. Å. S. Gustafson & J. C. Liou

Department of Astronomy, PO Box 112055, University of Florida, Gainesville, Florida 32611-2055, USA

Numerical simulations of the orbital evolution of asteroidal dust particles show that the Earth is embedded in a circumsolar ring of asteroidal dust, and has a cloud of dust permanently in its wake. This could account for the asymmetry of the zodiacal cloud observed by the Infrared Astronomical Satellite (IRAS). The resonant trapping and subsequent release of dust particles by the ring may provide a mechanism by which carbonaceous material is transported from the asteroid belt to the Earth.

ALL the observations of the distant Universe by spacecraft that operate in the infrared are made through a tenuous cloud of dust known as the zodiacal cloud that envelops the inner Solar System. Small, micrometre-sized particles in this cloud are removed by drag forces (Poynting–Robertson light drag and solar wind drag, that cause the particles to spiral in towards the sun)¹ on timescales $\sim 10^4$ years but are, presumably, continuously replenished². Some of these interplanetary dust particles are deposited in, and collected from, the Earth's upper atmosphere and constitute a major source of extraterrestrial material³. However, their origin, whether asteroidal or cometary, is a matter of dispute².

In 1983, it was discovered that the zodiacal cloud is not featureless. IRAS revealed circumsolar, near-ecliptic bands of dust⁴ that originate in the asteroid belt, but probably extend to inside the orbit of the Earth⁵, and seem to be related to the prominent asteroid families^{6,7}, that is, to known collision products in the asteroid belt. Modelling shows that the particles in these dust bands contribute $\sim 10\%$ of the thermal flux emitted by the zodiacal cloud and that dust from the main belt (between 2.2 and 3.3 astronomical units, where 1 AU is the Earth–Sun distance) is probably the source of about one-third of the thermal flux⁸. Due to the stochastic nature of collisions in the asteroid belt and our lack of knowledge of the number of small, collisionally-evolved asteroids in the outer regions of the asteroid belt, the latter estimate may be uncertain by a factor of two or more⁹. But we now know that the asteroid belt is certainly a significant source of zodiacal dust.

Gold¹⁰, in a prescient paper, pointed out that small particles in the Solar System that spiral in towards the Sun (because of Poynting–Robertson light drag) may become trapped in resonances with the planets, with the result that the zodiacal cloud may have a banded appearance when viewed from outside the central plane. Resonances occur at those heliocentric distances for which the ratio of the orbital periods of the particles and the planets are ratios of small integers. In recent years, orbital evolution and resonant trapping of dust grains have been discussed by numerous authors with regard to planetary accretion^{11–15} and the structure of the disk around β Pictoris^{16–18}. From their numerical investigations of the orbital evolution of small dust grains, Jackson and Zook¹⁹ made the specific suggestion that the Earth may be embedded in a ring of asteroidal particles in resonant lock with the planet.

The IRAS spacecraft has given us our highest resolution data on the structure of the zodiacal cloud. Dermott *et al.*²⁰ analysed these data and pointed out a marked but peculiar asymmetry, namely that the peak brightness of the zodiacal cloud in the

trailing direction (that is, opposite to the Earth's orbital motion—see Fig. 1) is consistently greater than that in the leading direction. This finding was confirmed by Reach²¹ who concluded, without detailed argument that either (1) there is a calibration inconsistency (in the gain) between the leading and trailing IRAS scans related to a radiation-induced responsivity enhancement produced by passage of the detectors through the South Atlantic Anomaly, or (2) there is an enhancement in dust density that follows the Earth around the Sun and that this could be related to the resonant trapping described by Jackson and Zook¹⁹ or could be due to gravitational focusing.

We do not argue here that the IRAS data are devoid of calibration problems (hysteresis^{21,22}, for example, that is associated with passage through the South Atlantic Anomaly is clearly visible in the 25 μm data shown in Fig. 6a) or that these calibration problems are fully understood. The calibration scheme was based on a simple model of the zodiacal cloud and, as our knowledge of the observed and modelled structure of the zodiacal cloud grows, the IRAS model will inevitably prove to be inadequate. Our purpose is to show (1) that if, as we believe, asteroidal collisions are a significant source of zodiacal dust, then a trailing–leading asymmetry of the zodiacal cloud is to be expected, and (2) that the asymmetry observed by IRAS is quantitatively consistent with predictions based on our numerical

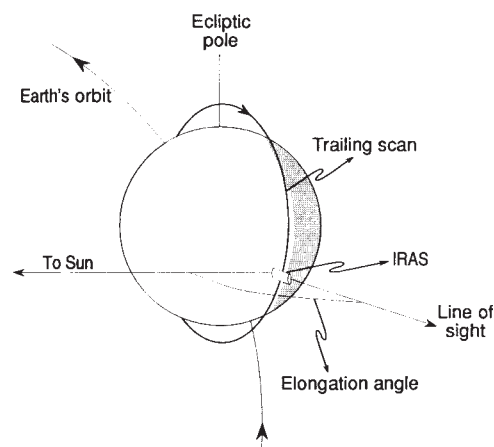


FIG. 1 Geometry of the IRAS observations. During each pole-to-pole scan of the sky, the elongation angle of the line of sight of the telescope was held constant.

investigations of the resonant trapping of asteroidal dust particles.

Dynamics of co-rotational resonance

The gravitational perturbations on the orbit of a dust particle in the Solar System can be separated into four distinct categories: short-period perturbations, resonant perturbations, secular (long-period) perturbations and direct scattering due to close encounters with the planets. The short-period perturbations are due to the gravitational attractions at planetary conjunction that repeat with the synodic period. The perturbations are resonant if the ratio of the periods of the bodies is a ratio of two small integers. Secular perturbations are due to the long-term averages of the forces between the planets and give rise to variations in the orbital eccentricities and inclinations. For particle diameters ranging from a few to a few hundred micrometres, non-gravitational forces are also important, particularly (1) light pressure and Poynting–Robertson light drag, and (2) solar wind drag due to the scattering of incident protons¹. Drag forces remove angular momentum from the particles causing them to spiral in towards the Sun on timescales that increase directly with their diameters and are $\sim 5 \times 10^4$ years for spherical, 12- μm diameter, astronomical silicate particles¹.

Until the terrestrial planets are encountered, secular perturbations and drag forces largely determine the orbital evolution. The problem of describing orbital evolution due to secular perturbations in the presence of non-gravitational forces has been solved for asteroidal particles. Dermott *et al.*⁵ have shown that the concepts of free and forced eccentricities and inclinations that arise from the solution of the secular perturbation equations without drag remain valid in the presence of drag. In particular, the proper inclinations remain constant whereas the proper eccentricities, e_0 , decay with the semimajor axis, a , according to the two-body formulation of Wyatt and Whipple²³, that is,

$$\frac{de_0}{da} = \frac{5}{2} \frac{e_0}{a} \frac{(1 - e_0^2)}{(2 + 3e_0^2)} \quad (1)$$

and is independent of particle size. The forced inclinations and eccentricities can be calculated once the drag rates are specified⁵.

Strong resonant interactions come into play as the particles spiral through the asteroid belt and encounter the Kirkwood gaps that correspond to mean motion resonances with Jupiter. Because the particles are moving away from Jupiter, that is, because their orbits and that of Jupiter are diverging, capture into resonance is unlikely and passage through these jovian resonances has little influence on the orbital elements^{12,24}. But the same particles move on orbits that converge on those of the terrestrial planets and, although the terrestrial resonances are weaker than the jovian resonances, evolution on converging orbits allows capture into resonance²⁴. We confirm this by direct numerical integration of the full equations of motion for 12- μm -diameter asteroidal particles using the RADAU fifteenth-order integrator program with variable time steps taken at Gauss–Radau spacings²⁵. We consider spherical astronomical silicate¹ particles of density 2.7 g cm^{-3} for which the ratio of radiation pressure to gravitational force, β , is 0.037. The average force due to the solar wind is taken to be 30% of the Poynting–Robertson light drag force, varying on a 11-year cycle from 20% to 40% (ref. 1). The initial distributions of the proper inclinations and eccentricities are the same as those of the brighter asteroids in the main belt. Under these conditions, numerical integration of the orbits of 912 particles originating in the main belt shows that $\sim 20\%$ of these particles are trapped in first-order co-rotational resonances outside the Earth's orbit (Fig. 2).

If we retain only the leading term in the disturbing function of the particle, then the equation of motion of the resonant argument, ϕ , defined by

$$\phi = p\lambda - (p+1)\lambda' + \bar{\omega}' \quad (2)$$

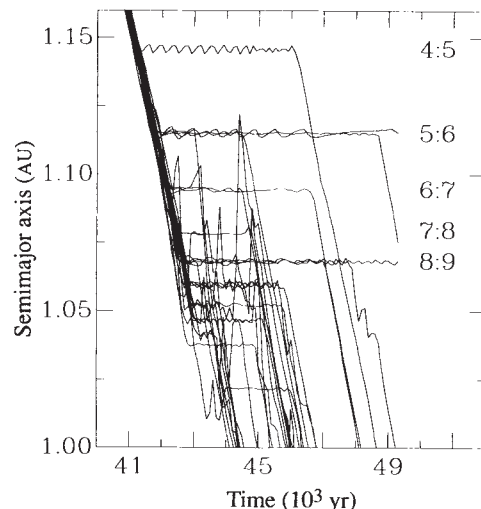


FIG. 2 Orbital evolution of 12- μm -diameter asteroidal dust particles of density 2.7 g cm^{-3} and $\beta = 0.037$, showing capture into (and escape from) $p:(p+1)$ resonances outside the orbit of the Earth (see text and equation (2) for definition of β and p). The orbits of 912 particles with initial elements identical to those of a bias-free set of main-belt asteroids were integrated numerically. This plot shows the orbital history of the 28 of the 114 particles originating in the inner part of the belt that were trapped in resonance with the Earth.

is that of a damped harmonic oscillator and the acceleration of ϕ , $\ddot{\phi}$, is given by

$$\ddot{\phi} = -(Gm/a')f(\alpha, \beta)e' \sin \phi - (p+1)\dot{n}'_a \quad (3)$$

where p is an integer, λ is the mean longitude, $\bar{\omega}$ is the longitude of perihelion, Gm is the gravitational mass of the Earth, a is the semimajor axis, e the eccentricity, $\alpha = a/a'$, $f(\alpha, \beta)$ is a function of Laplace coefficients and β that increases markedly with increasing p , $\dot{n}'_a$ is the average rate of change of the mean motion n' due to the action of drag, unprimed quantities refer to the orbit of the Earth and primed quantities refer to the orbit of the dust particle¹³.

The path of a particle librating in a 5:6 corotational resonance with the Earth is shown in Fig. 3. Without drag, the path would have mirror symmetry about the mean Earth–Sun line. But drag introduces a phase lag into the equation of motion, with the result that the path is asymmetric; as the particles pass through perihelion, they approach the Earth closer in the trailing direction (behind the Earth in its orbit) than in the leading direction.

Predicted ring structure

By setting the left-hand side of equation (3) to zero, we find that the phase lag is a maximum ($|\sin \phi| \approx 1$), when

$$|(Gm/a')f(\alpha, \beta)e'| \approx |(p+1)\dot{n}'_a| \quad (4)$$

and the strength of the resonance is just sufficient to counteract the effect of drag on the particle's semimajor axis. While the particle is trapped in resonance, drag acts quickly to increase the particle's eccentricity e' and the strength of the resonance²⁴. Paradoxically, this leads to resonance disruption on timescales $\lesssim 10^4$ years (Figs 2 and 4). This is caused by the onset of chaos due to resonance overlap²⁴ and, probably more importantly, direct particle scattering due to close encounters with the Earth^{26–28}, once drag has reduced the particle's perihelion to inside the Earth's orbit. The disruption timescales are smaller than, but comparable with, the collision lifetimes of the particles², and interparticle collisions will have to be considered in a more complete analysis of the ring's structure.

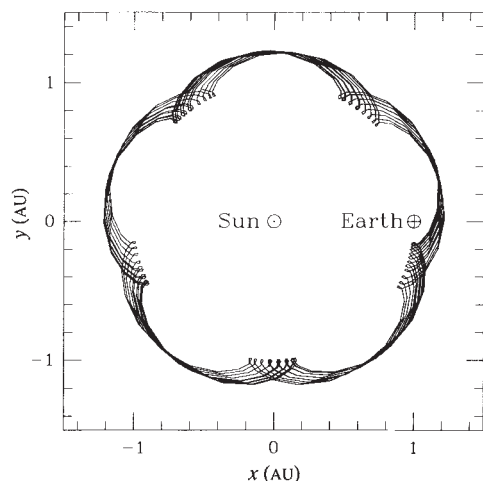


FIG. 3 The path of a 12- μ m-diameter asteroidal dust particle trapped in a 5:6 resonance with the Earth is shown, not in an inertial frame, but in a frame centred on the Sun and co-rotating with the Earth's mean motion (the Earth is near-stationary in this frame). The motion of the particle is tracked over one complete libration period of the resonant argument. Particles of this size experience large drag forces, with the result that their paths are not symmetric about the Sun–Earth line.

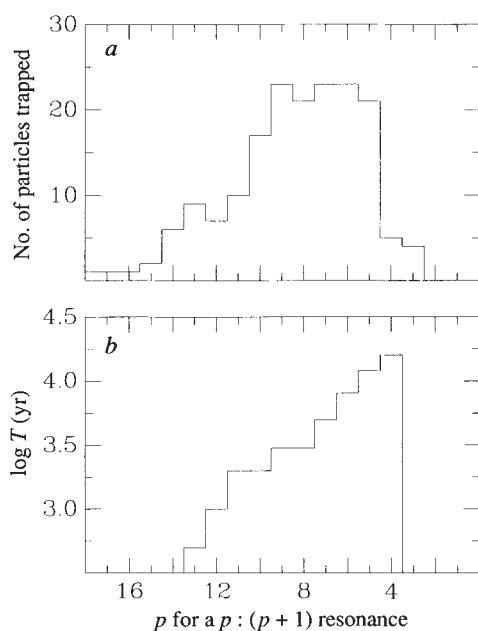


FIG. 4 *a*, The relative probability of capture into, and *b* the average trapping time T in, $p:(p+1)$ resonances outside the orbit of the Earth of 12- μ m-diameter asteroidal dust particles. The orbits of 912 particles with initial elements identical to those of a bias-free set of main-belt asteroids were integrated numerically. The p values of the 169 particles that were trapped in resonance are shown in *a*, and their average trapping times in those resonances are shown in *b*.

The structure of the ring produced by resonant trapping is estimated from our numerical integrations. We consider main-belt asteroidal particles, because the high orbital eccentricities of both cometary particles and particles deriving from the disruption of near-Earth asteroids make resonant trapping highly improbable^{19,29}. The size–frequency distribution of asteroidal particles is such that the effective area of dust, and hence the thermal flux from the ring observed in a given IRAS waveband, may be dominated by flux from the near-smallest particles in the distribution (that is, particles with radii $\gtrsim \Lambda/2\pi$ —where Λ

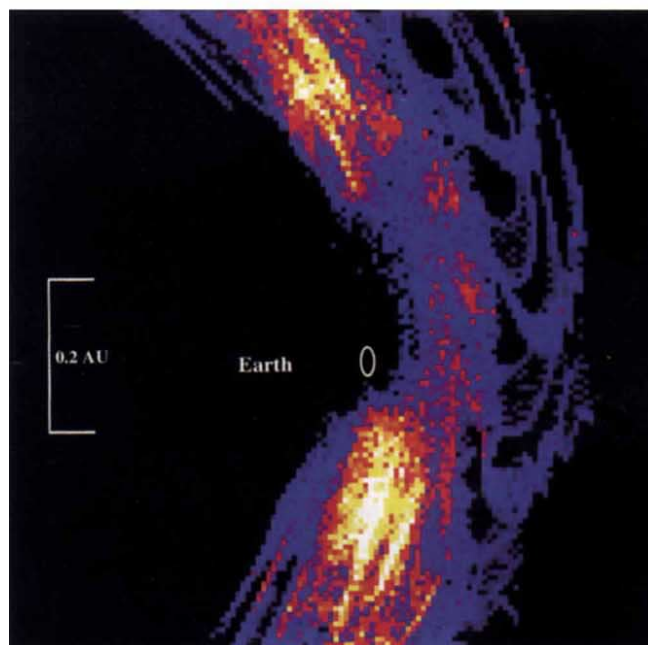
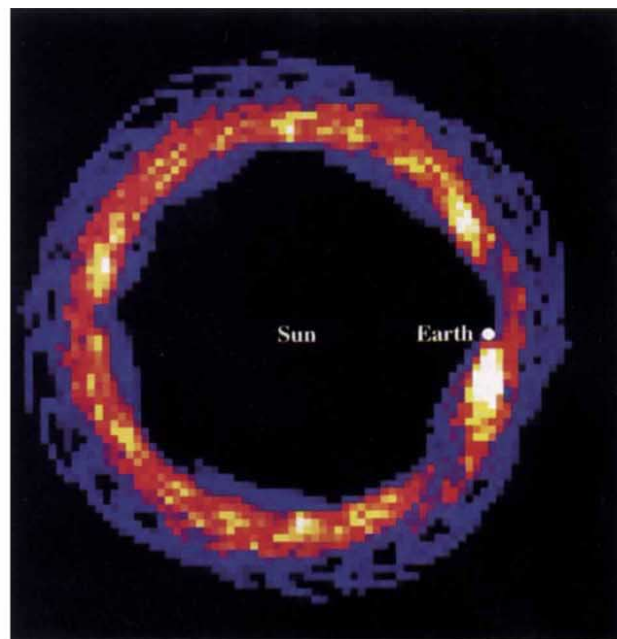


FIG. 5 Top, numerical simulation of the structure of a heliocentric ring of asteroidal dust particles in a frame centred on the Sun and co-rotating with the Earth. The cloud of particles that appears to trail the Earth in its orbit has a peak particle number density (colour-coded white) $\sim 10\%$ greater than that of the background zodiacal cloud. The resolution of the image is (0.04×0.04) AU. Bottom, higher-resolution image (0.01×0.01) AU of the numerical simulation shown above. On this scale, the eccentricity of the Earth's orbit ($e=0.0167$) is apparent and the Earth's path in the co-rotating frame is depicted by the small 2:1 ellipse: motion around this ellipse is clockwise. The Earth passes through perihelion in the first week of January, and is closest to the trailing part of the ring in the first week of October.

is the waveband of the detector—that are also large enough to satisfy the resonance strength criterion given by equation (4)). These particles also have the highest phase shifts, with the result that the dynamics of resonance is in the non-adiabatic regime²⁴ (for which analytic theory is only a guide) and reliable results must be obtained by numerical integration. The drag rates on very small particles can be too high to allow resonance trapping,

the critical size varying with the particle's eccentricity e' and the p value of the resonance²⁶. To obtain a detailed description of the structure of the ring, the dynamics of a wide range of particle sizes need to be considered. The size-frequency distribution of asteroidal particles near the Earth is uncertain, especially if asteroidal particles are not the dominant source of the zodiacal cloud. We estimate that asteroidal particles of diameter 12 μm and density 2.7 g cm^{-3} are close to the critical size for trapping into the most significant resonances. We have also shown that the dynamics of particles of about this size can account for the observed plane of symmetry of the background zodiacal cloud⁵. Here we estimate the structure of the ring from the dynamics of these particles alone.

We consider that a particle is trapped if it remains in a given resonance for at least ~ 500 years, or more than one libration period. On integrating the orbits of 912 particles, we find that up to 20% of the particles are trapped in first-order resonances with $p:(p+1)$ ranging from 3:4 to 17:18 with occasionally even higher values of p . Figure 4a shows the relative probability of capture into the individual resonances. We determine the average trapping time for particles in these resonances empirically (Fig. 4b). The trapping times decrease sharply with increasing p as the locations of the resonances get closer to the Earth and the probability of close encounters increases. The overall distribution of particles in the ring at any given time is a convolution of the quantities given in Fig. 4a and b with the distributions of the positions of the particles in the various resonances, as tracked in the rotating reference frame, from the time of capture to the typical time of release.

Our final model is a simulated 'image' binned in pixels of 0.04×0.04 AU (Fig. 5). The 5:6 resonance shown in Fig. 3 has five lobes, with the Earth residing asymmetrically in one of those lobes. All other resonances have similar structures, only the number of lobes varies. The resultant image obtained by superposition of the various paths weighted according to the probability of capture and the trapping times shows that, in a rotating reference frame, the trapped particles form a near-uniform ring around the Sun that co-rotates in inertial space with the Earth.

The peripheral arc-like structures in the ring are artefacts caused by the limited number of single-size particles considered in the model. We consider, however, that the other noticeable features in Fig. 5 are real. First, the existence of a cavity in the ring at the location of the Earth (see also ref. 30). Figure 5 also shows that apart from an asymmetry in the mean position of the Earth in the ring's cavity, there is also a marked asymmetry in the longitudinal variation of the particle number density. The

increase in resonance strength—and the corresponding decrease in phase shift with increasing p value—disperses the longitudes of the perihelia of the particle paths in the rotating frame in the leading direction, while concentrating the longitudes of the perihelia in the trailing direction. This results in a marked enhancement of particle number density behind the Earth in its orbit, as if the Earth had a trailing cloud of dust permanently in its wake.

Observed zodiacal cloud asymmetry

The observed trailing-leading asymmetry is most clearly revealed by exploiting the IRAS observing sequence. To obtain good all-sky coverage, the elongation angle used by IRAS was not held constant at 90° , but was shifted systematically from scan to scan over periods of a few weeks to ensure complete coverage of the sky. For the first two-thirds of the mission, the elongation angle, while being kept within about 10° of 90° , was incrementally increased from scan to scan for scans in the trailing direction, whereas for scans in the leading direction the elongation was incrementally decreased. These regularities allow us to increase the signal-to-noise ratio by using all the data obtained in a given period (of a few weeks) to obtain (1) the average brightness for an elongation angle of 90° , and (2) the Earth's ecliptic longitude when that elongation angle was measured.

A plot of the variation with elongation angle of the peak near-ecliptic brightness of the broad-scale zodiacal background observed by IRAS in three wavebands from 9 to 11 July 1983 is shown in Fig. 6a (Fourier methods are used to separate the smooth, large-scale zodiacal background from the signals associated with the narrower asteroidal dust bands²⁰). We have obtained similar plots for all the other periods of the mission for which data were obtained that encompass an elongation angle of 90° and that are not overly contaminated by noise deriving from proximity to the Galactic plane. Variations of the average brightness for an elongation angle of 90° with the ecliptic longitude of the Earth in both the trailing and leading directions are shown in Fig. 6b.

There are three reasons why the peak brightness of the zodiacal cloud should vary with ecliptic longitude when viewed at a constant elongation angle³¹. These arise from the forced eccentricities of the dust particle orbits²⁰, from the inclination of the plane of symmetry of the cloud with respect to the ecliptic and from the Earth's orbital eccentricity. By modelling a cloud of asteroidal dust particles³¹, we have shown (1) that the Earth's orbital eccentricity is the dominant effect and (2) that the peak

FIG. 6 a, (Left-hand column) variation of the peak (near-ecliptic) brightness of the large-scale background zodiacal cloud with elongation angle, in three separate wavebands. The data in the leading direction (that is, in the direction of the Earth's orbital motion) are from 9 to 17 July 1983 (open circles). The data in the trailing direction (that is, in the direction opposite to the Earth's orbital motion) are from 4 to 12 July 1983 (filled circles). The best-fit straight lines are used to determine the peak fluxes corresponding to an elongation angle of 90° . b, (right-hand column) variation with the Earth's ecliptic longitude (or time of the year) of the peak (near-ecliptic) brightness of the large-scale zodiacal cloud observed by IRAS in three separate wavebands at an elongation angle of 90° in the trailing direction (solid line and filled circles) and in the leading direction (broken line and open circles). In all cases except one, the magnitudes of the error bars of the points (derived from plots similar to those in a) are too small to show.

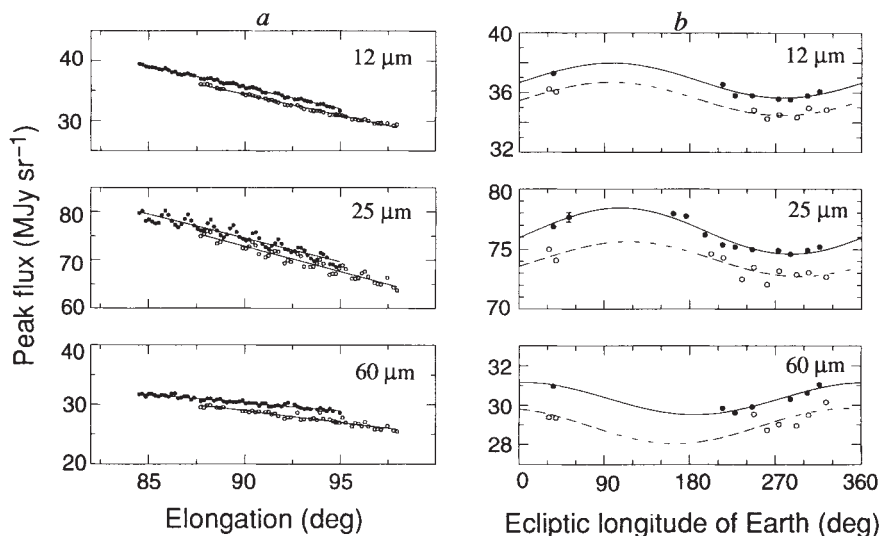


TABLE 1 Variation of the peak flux observed by IRAS

Waveband μm	Direction	Longitude of Earth for peak flux (degrees)	Mean flux (MJy sr^{-1})	Flux ratio (T/L)
12	T	98 ± 6	36.82 ± 0.13	1.035 ± 0.007
12	L	97 ± 13	35.59 ± 0.21	
25	T	106 ± 5	76.52 ± 0.16	1.031 ± 0.004
25	L	118 ± 17	74.19 ± 0.26	
60	T	5 ± 11	30.34 ± 0.13	1.045 ± 0.016
60	L	345 ± 26	29.03 ± 0.44	

These measurements were made at an elongation angle of 90° in the trailing (T) and leading (L) directions.

brightness in both the leading and trailing directions should occur when the Earth is close to perihelion, that is, at an ecliptic longitude of 102.3° . These variations are clearly observed in both the 12- μm and the 25- μm wavebands (Fig. 6b and Table 1). The weighted average of the ecliptic longitudes of the Earth associated with the peaks of the sine curves fitted to the 12- μm and 25- μm data is $103 \pm 5^\circ$. The data in the 60- μm waveband are not consistent with those in the other wavebands. This could be due to a contribution of thermal flux from the Galaxy, but this needs to be verified. But ratios of fluxes are insensitive to calibration errors, and Fig. 6b shows that in all three wavebands, at all times of the year, observations in the trailing direction are brighter than those in the leading direction by 3 or 4% (the precise mean ratios for each waveband, that depend on fitting sine curves through the data, are given in Table 1).

We consider that the asymmetries of the Earth's resonant ring that arise from (1) the position of the Earth near the trailing edge of a cavity in the ring and (2) the cloud of asteroidal particles that trails the Earth in its orbit, act together to produce the observed trailing-to-leading flux ratios. If we assume that all the particles in the zodiacal cloud are asteroidal and of diameter 12 μm , then we calculate that the asymmetry between the trailing and leading flux due to the ring is as high as 12%. But recent calculations, using two new results (that the particles in the dust bands are asteroidal and constitute 10% of the cloud, and that the ratio of the thermal flux from the asteroidal families to that from the entire main-belt of asteroids is 1:3.4), lead us to estimate that about one-third of the zodiacal cloud is asteroidal in origin⁸ (see also Reach³²). Therefore, if only one-third of the cloud is made up of asteroidal particles (that is, particles of low orbital eccentricity) that can be trapped in the ring, the asymmetry in the flux will decrease from 12 to 4% which is comparable to the observed value. But the dust in the zodiacal cloud must have a size-frequency distribution and its effect has to be studied carefully. Because asteroidal particles in the cloud with diameters

$\lesssim 5 \mu\text{m}$ do not get trapped in the ring, they will tend to decrease the amount of asymmetry, whereas the larger particles in the cloud, although less in number, have higher probabilities of capture into resonances (because of their smaller drag rates) and will increase the flux from the ring. Using data from the Cosmic Background Explorer (COBE) satellite, we will obtain a more precise estimate of the trailing-leading asymmetry, not only at 25 μm but at a range of wavelengths which will enable us to develop a more sophisticated model of the ring that includes a range of particle sizes. Precise measurements of the ratio of trailing flux to leading flux will allow us to determine the contribution of asteroidal particles to the near-Earth region of the zodiacal cloud, and may place constraints on their size-frequency distribution.

From the model shown in Fig. 5, we can predict two effects due to the Earth's epicycle: (1) a yearly variation in the flux ratio, and (2) a higher probability of interplanetary dust particles encountering the Earth at certain times of the year. Both of these effects will be a maximum in the second half of the year (around September/October) when the Earth is closest to the trailing cloud. This annual variation of the asymmetry will provide further constraints on the number distribution of asteroidal particles in the zodiacal cloud, and can be determined from the COBE data which has more accurate flux calibration as well as extensive coverage in both elongation angle and wavelength. The release of particles from the ring due to close encounters may make the ring act as a funnel which directs particles in near-circular orbits into the Earth's upper atmosphere. This may prove to be a mechanism for the delivery of carbonaceous material from the asteroid belt to the Earth. Further studies are needed to determine the fraction of particles actually colliding with the Earth.

This mechanism of ring formation is not exclusive to the Earth and can be extended to other terrestrial planets. Both Mars and Venus will favour the capture of larger particles which have lower drag-rates. This is because even though the particles have lower drag rates near Mars, the martian mass is only 10% of that of the Earth's and the martian resonances are correspondingly weaker. In the case of Venus, the particles are moving with a higher drag-rate. Mercury is unable to capture significant numbers of particles because of its very low mass and the higher drag-rates of particles close to the Sun.

Extending this phenomenon outside our Solar System, the discovery of large disks around stars like β Pictoris (by IRAS) and Fomalhaut³³ suggests that resonance effects may play an important role in the detection of other planetary systems. Resonances can give rise to features like arcs and gaps in the disk of dust surrounding the star, indicating the presence of a planet which may otherwise be undetectable¹⁵⁻¹⁸. □

Received 8 April; accepted 25 May 1994.

- Gustafson, B. Å. *S. A. Rev. Earth & planet. Sci.* **22**, 553-595 (1994).
- Grün, E., Zook, H. A., Fechtig, H. & Giese, R. H. *Icarus* **62**, 244-272 (1985).
- Bradley, J. P., Sandford, S. A. & Walker, R. M. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews M. S.) 861-895 (Univ. Arizona, Tucson, 1988).
- Low, F. J. et al. *Astrophys. J.* **278**, L19-L22 (1984).
- Dermott, S. F. et al. in *Chaos, Resonance, and Collective Dynamical Phenomena in the Solar System* (ed. Ferraz-Mello, S.) 333-347 (Kluwer, Dordrecht, 1992).
- Dermott, S. F., Nicholson, P. D., Burns, J. A. & Houck, J. R. *Nature* **312**, 505-509 (1984).
- Sykes, M. V. *Icarus* **84**, 267-289 (1990).
- Dermott, S. F. et al. in *Asteroids, Comets, Meteors 1993* (eds Milani, A. & Di Martino, M.) (Kluwer, Dordrecht, in the press).
- Durda, D. D. thesis, Univ. Florida (1993).
- Gold, T. *Icarus* **25**, 489-491 (1975).
- Greenberg, R. *Icarus* **33**, 62-73 (1978).
- Gonczi, R., Froeschlé, Ch. & Froeschlé, C. *Icarus* **51**, 633-654 (1982).
- Weidenschilling, D. & Davis, D. *Icarus* **62**, 16-29 (1985).
- Patterson, C. W. *Icarus* **70**, 319-333 (1987).
- Sicardy, B., Beaugé, C., Ferraz-Mello, S., Lazzaro, D. & Roques, F. *Celest. Mech. dyn. Astr.* **57**, 373-390 (1993).
- Scholl, H., Roques, F. & Sicardy, B. *Celest. Mech. dyn. Astr.* **56**, 381-393 (1993).
- Roques, F., Scholl, H., Sicardy, B. & Smith, B. *Icarus* **108**, 37-58 (1994).
- Lazzaro, D., Sicardy, B., Roques, F. & Greenberg, R. *Icarus* **108**, 59-80 (1994).

- Jackson, A. A. & Zook, H. A. *Nature* **337**, 629-631 (1989).
- Dermott, S. F., Nicholson, P. D., Kim, Y., Wolven, B. & Tedesco, E. F. in *Comets to Cosmology* (ed. Lawrence, A.) 3-18 (Springer, Berlin, 1988).
- Reach, W. T. *Astrophys. J.* **369**, 529-543 (1991).
- Deul, E. R. & Walker, H. J. *Astr. Astrophys. Suppl. Ser.* **81**, 207-214 (1989).
- Wyatt, S. P. & Whipple, F. L. *Astrophys. J.* **111**, 134-141 (1950).
- Dermott, S. F., Malhotra, R. & Murray, C. D. *Icarus* **76**, 295-334 (1988).
- Everhart, E. in *Dynamics of Comets* (eds Carusi, A. & Valsecchi, G. B.) 185-202 (Reidel, Dordrecht, 1985).
- Weidenschilling, S. & Jackson, A. A. *Icarus* **104**, 244-254 (1993).
- Jackson, A. A. & Zook, H. A. *Icarus* **97**, 70-84 (1992).
- Marzari, F. & Vanzani, V. *Astr. Astrophys.* **283**, 275-286 (1994).
- Liou, J. C. thesis, Univ. Florida (1993).
- Jackson, A. A. & Zook, H. A. *Lunar planet. Sci.* **XXIII**, 595-596 (1992).
- Xu, Y. L. et al. in *The Ceremony of the 70th Anniversary of the Chinese Astronomical Society* (Chinese Acad. Sci., Beijing, in the press).
- Reach, W. T. in *Origin and Evolution of Interplanetary Dust* (eds Levasseur, A. C. & Hasegawa, H.) 211-214 (Kluwer, Dordrecht, 1991).
- Stern, S. A., Weintraub, D. A. & Festou, M. C. *Nature* **368**, 312-314 (1994).

ACKNOWLEDGEMENTS. We thank J. A. Burns and W. T. Reach for comments, and the Department of Computer Sciences at the University of Florida for use of the KSR1 Parallel Supercomputer. This work was supported by NASA, the University of Florida, the IBM Corporation and the Geophysics Directorate of the US Air Force Phillips Laboratory.

The role of non-deformable units in pressure-induced reversible amorphization of clathrasils

J. S. Tse*, D. D. Klug*, J. A. Ripmeester*, S. Desgreniers† & K. Lagarec†

* Steacie Institute for Molecular Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

† Department of Physics, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada

PRESSURE-induced amorphization of solids has been much studied since it was first observed in 1984¹. It was found recently^{2,3} that some materials can be amorphized reversibly under pressure, reverting back to the original crystalline structure and orientation when the pressure is decreased. It has been suggested⁴ that the presence of non-deformable units is essential for this reversibility, these units acting as templates around which the original structure is reformed. Here we investigate this idea by comparing the effect of pressure on two clathrasils—silica solids with open, microporous structures—with and without guest molecules inside the pores. We have studied pressure-induced amorphization of dodecasil-3C, which has cage-like voids, and dodecasil-3R, which has two-dimensional channels. For both materials, our experiments and simulations show that amorphization is fully or partly reversible only when guest molecules are present, suggesting that these do indeed act as rigid ‘organizing centres’ for the reversible transformation.

Certain crystalline clathrate hydrates can be compressed into a higher-density amorphous phase under high pressure^{2,5}, but when the pressure is removed, X-ray diffraction indicates that the recovered sample corresponds to the original crystalline clathrate⁵. The pressure-amorphized phase of ice Ih, on the other hand, can be preserved at 77 K under ambient pressure¹. In view of the similarity in the short-range order of water molecules in both ice and clathrate hydrates⁶, it is surprising that the densified, amorphous clathrate phases are unstable towards recrystallization at ambient pressure. This raises the question of whether the guest molecules present in the clathrate influence the stability of the compressed structure. Although in principle one might test this conjecture by subjecting a guest-free hydrate to pressurization, clathrate hydrates without the presence of guests are thermodynamically unstable⁷. As an alternative, we have therefore considered here a class of inclusion compounds with characteristics closely related to clathrate hydrates but for which both filled and empty structures are stable⁸. Clathrasil dodecasil-3C (D3C), a new class of porous tectosilicate, is structurally isomorphic with type II hydrates⁹ and has cage-like voids in which guest species can be enclathrated. The empty clathrate can be prepared by thermal decomposition of the organic guest molecules. More significantly, the common crystalline silicate, α -quartz, also behaves like ice Ih where the pressure-amorphized phase is recoverable at ambient pressure¹⁰.

To define the role of the host lattice, it is necessary to make a comparative study of the effect of pressure on a related compound of different structure. Clathrasil dodecasil-3R (DD3R) was chosen for this purpose. Unlike the closed cages in D3C, the DD3R cavities are much larger and are interconnected (forming a two-dimensional pore system by sharing the faces of SiO₂ polygons), allowing the diffusive exchange of small guest molecules¹¹.

Samples of dodecasil-3C with enclathrated tetrahydrofuran (D3C/THF) and dodecasil-3R with amino-adamantane (DD3R/ADM) were synthesized according to the published procedure¹¹. The empty DD3R was prepared in a similar way

in the absence of the guest. The empty D3C structure was obtained by calcination of a D3C/THF sample at 900 °C for 24 hours. The infrared spectra of D3C/THF and the calcined sample as a function of pressure are shown in Fig. 1. For the D3C/THF sample, the following spectral changes were observed with increasing pressure. There are three main features in the spectrum. The relatively sharp O–Si–O bending mode^{12,13} around 500 cm^{−1} decreases in intensity, shifts to slightly higher frequency, abruptly disappearing at ~4.5 GPa. The apparent disappearance of this band is associated with the phase transformation observed in X-ray diffraction (see below). Close to the phase transition, the mixed Si–O–Si bending and Si–O stretching^{12,13} band centred at 800 cm^{−1} shifts to higher frequency and broadens. The predominantly Si–O stretching band at 1,050 cm^{−1} remains largely unchanged. The behaviour of these vibrational modes are consistent with the O–Si–O linkages being the most affected at high pressure. The blue-shift in both the O–Si–O bending and the 800 cm^{−1} Si–O stretching bands are consistent with more compressed SiO₂ units where the O–Si–O angles are distorted. These structural changes agree well with those observed in pressurized α -quartz¹⁴ and silica glass¹⁵. Similar spectral changes were also observed in compressed ice¹⁶.

There is a remarkable change in the infrared spectrum of D3C/THF when the pressure is lowered from 11.7 GPa: the sharp feature at 500 cm^{−1} reappears. Moreover, the profiles and positions of the other vibrational bands are identical to those of the starting material. These observations indicate strongly that

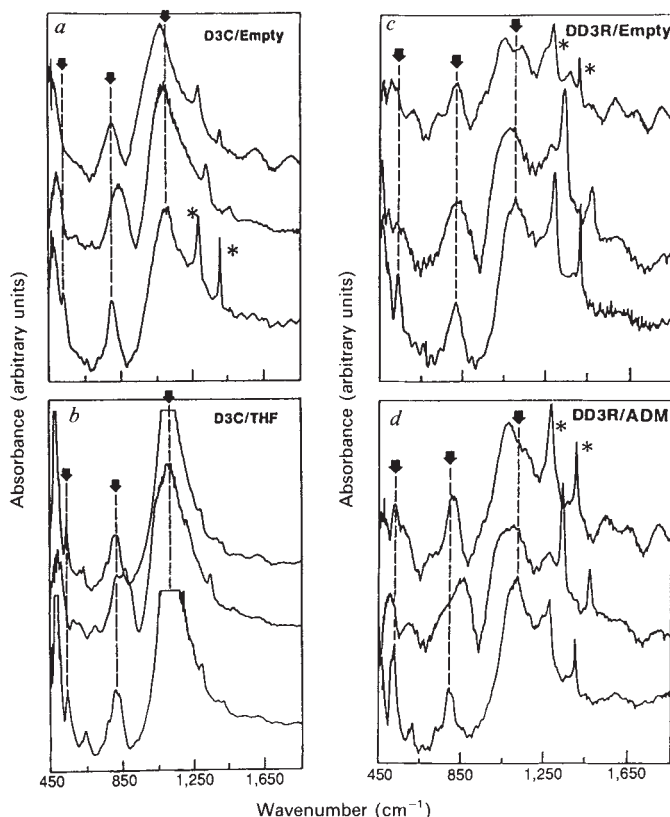


FIG. 1 Infrared spectra of clathrasils measured in a diamond anvil with dilute solid solutions of sodium nitrite and nitrate in sodium bromide as pressure calibrants³¹. **a**, Empty D3C at 1 GPa (bottom trace), 12.8 GPa (middle) and decreased to 0.7 GPa after transformation (top). **b**, D3C/THF at 2 GPa (bottom), 11.7 GPa (middle) and after decreasing to 0.8 GPa after transformation (top). **c**, Empty DD3R at 1 GPa (bottom), 10.8 GPa (middle) and decreased to 0.2 GPa after transformation (top). **d**, DD3R/ADM at 1 GPa (bottom), 10.3 GPa (middle) and decreased to 0.2 GPa after transformation (top). The peaks marked with an asterisk are due to the internal pressure standard NaNO₂/NaNO₃. Arrows indicate the infrared peak positions at ambient pressure.

the compressed phase has reverted back to the original crystalline structure.

The empty D3C shows similar spectral changes with increasing pressure. The apparent transition pressure for D3C is again at ~ 4.5 GPa. However, when the pressure is decreased from 12.8 GPa, there is no significant change in the infrared profile as was observed in D3C/THF. The only spectral changes are the expected frequency shifts with decreasing pressure. Therefore, the empty D3C seems to retain the compressed framework at ambient pressure. As the lattice framework for both D3C and D3C/THF are practically the same, the reversible transformation in the latter must be attributed to the ubiquitous presence of the guest molecules.

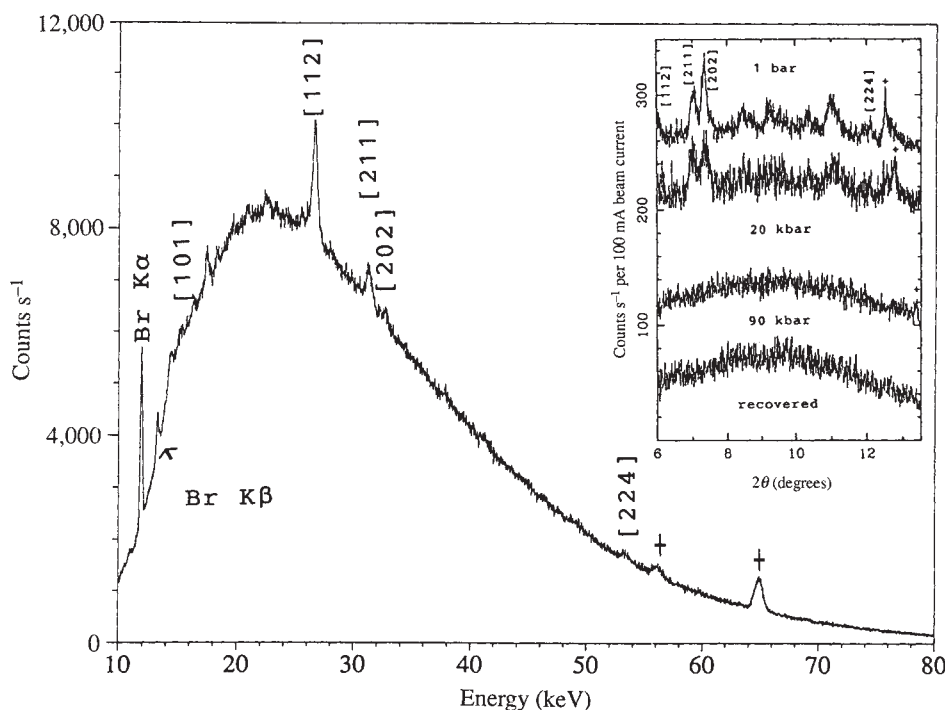
The deductions made from the infrared studies are corroborated by the X-ray diffraction studies. The Bragg reflections of the crystalline tetragonal structure¹⁷ of D3C/THF at room temperature can be clearly identified, and remain unchanged up to a pressure of ~ 6 GPa. Beyond that, the diffraction peaks disappeared and the pattern resembles that of an amorphous material. The sample did not revert immediately to the crystalline form on release of pressure as in the infrared experiment. However, there are strong indications of crystallinity when the sample is allowed to relax under ambient pressure for 12 hours. Another compressed sample was prepared under similar pressure conditions and was allowed to relax for 2 weeks. The energy dispersive diffraction pattern in Fig. 2 clearly show that the Bragg reflections of the original crystalline phase have reappeared. The difference in the relaxation behaviour of compressed D3C/THF as determined by infrared and diffraction experiments is probably due to the fact that the techniques reflect short-range and long-range order respectively, and it is not unusual for the two relaxation processes to proceed at different rates. Furthermore, the difference in pressure media may also contribute. In the study of pressure-induced amorphization of clathrate hydrates, it was found that the reversion process is also strongly kinetically controlled². The empty D3C sample shows a cubic pattern at ambient pressure¹⁷. Amorphization occurs at ~ 6 GPa, with no hint of crystallinity observed in the recovered sample.

The infrared spectra for DD3R/ADM and empty DD3R are shown in Fig. 1. The variation of infrared bands as a function

of pressure is very similar to that in D3C. The empty DD3R was found to transform to a (probably) disordered phase at 4 GPa and apparently, the high-pressure structure is retained when recovered at ambient pressure. An abrupt change in spectral features was observed in DD3R/ADM around 7 GPa, which indicates a structural transformation. At first glance, it appears that the infrared spectrum of the transformed sample after the pressure is relieved is broadly similar to that of the original sample. However, there are subtle differences in detail. In particular, the band at 700 cm^{-1} does not recover its shape, and the predominantly Si-O stretching band at $1,050\text{ cm}^{-1}$ is $\sim 40\text{ cm}^{-1}$ lower than in the starting material. On the other hand, the O-Si-O (500 cm^{-1}) and the Si-O (800 cm^{-1}) stretching modes apparently retain the same vibrational frequencies as the starting material, but the linewidth is significantly broader. These small differences are significant, as they probably indicate that the compressed sample has not fully reverted to the crystalline structure when the pressure is removed.

Recent experiments^{2,18} and molecular dynamics calculations^{19,20} indicate that a mechanical instability^{21,22} in the structure is the cause for pressure-induced amorphization. When a crystalline solid is subjected to external pressure, it compresses elastically until a mechanical instability is reached, and then collapses into a disordered structure^{19,20,23}. In clathrasils and clathrate hydrates at low temperature where atomic diffusion is restricted, the host lattice molecules can only distort around the guests and, therefore, retain some fundamental short-range structural features. Because the guests are covalently bonded entities, their molecular structures are little affected by moderate pressure. The absence of Bragg reflections in the X-ray powder pattern diffraction of the transformed phase only indicates the loss of long-range order. When the external stress is removed, the repulsive guest-host interactions act as centres for directing the host molecules back to their original positions. For open structures such as ice Ih α -quartz and the empty clathrasil, the lattices collapse into the voids^{19,20,23} and there are no directional forces, corresponding to those found in the inclusion compounds, to help restore the crystalline structure when the pressure is removed². The absence of repulsive interactions in the open channels helps to explain the partial reversion to the crystalline structure in pressure-amorphized DD3R/ADM. In

FIG. 2 Angle dispersive diffraction pattern for a sample of D3C/THF on compression (inset) and the energy dispersive diffraction pattern of a sample recovered at ambient pressure after being compressed to 10.5 GPa and allowed to relax for 2 weeks (main Figure). Peaks marked with + denotes reflection from NaCl, which acts both as pressure transmission medium and pressure calibrant³². Note the presence of small amount of impurities in the region 18–26 keV. X-ray diffraction experiments were performed at the National Synchrotron Light Source (NSLS) and at Cornell High Energy Synchrotron Source (CHESS). Angle dispersive diffraction was performed on beamline X-7A at NSLS using radiation of $0.7028\text{ \AA}$, monochromatized with a Ge [111] crystal³³. A position-sensitive detector was used in data collection³⁴. Energy dispersive diffraction was performed at beamline B1 of CHESS using white synchrotron radiation collimated to $150\text{ }\mu\text{m}$ and an intrinsic Ge detector set at a diffracting angle of $2\theta = 3.88^\circ$. A modified Merrill-Bassett diamond anvil cell³⁵, with a $25\text{--}50\text{ }\mu\text{m}$ thick sample supported in a stainless-steel gasket with a hole diameter of $250\text{ }\mu\text{m}$, was used in both experiments. Arrow indicates fluorescence from Br impurity in NaCl.



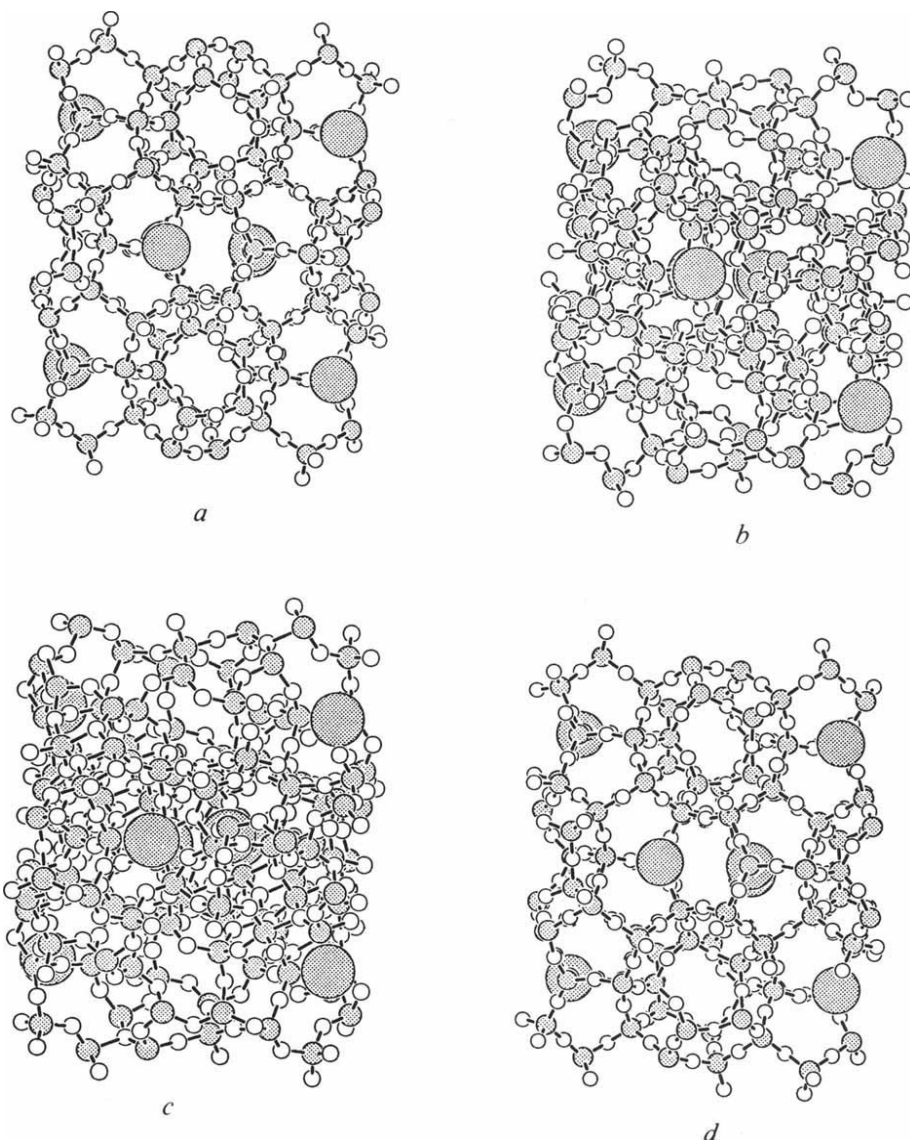


FIG. 3 Framework structures for a model D3C compound with eight spherical guests, as predicted by a constant-pressure, constant-enthalpy molecular dynamics calculation. The cell used in the calculations contains 136 SiO_2 units (Si atoms, small shaded spheres; O atoms, unshaded spheres). The intermolecular potentials for SiO_2 are taken from van Beest et al.³⁶. A Buckingham potential ($Ae^{-Br} - C/r^6$) was chosen for the spherical guest X (large shaded spheres) with parameters that mimic the interactions between two THF molecules. The parameters are $A = 7,370 \text{ kcal mol}^{-1}$, $B = 1.71 \text{ \AA}$ and $C = 16,281 \text{ kcal mol}^{-1}$. The usual combination rules were employed to construct the potential parameters for the Si-X and O-X interactions. Pressures: 1 bar (a), 10 GPa (b), compressed to 15 GPa and the pressure gradually lifted to 5 GPa (c), and sample recovered at 1 bar after compression (d).

this case, only the SiO_2 units in the vicinity of the guests are expected to revert back to their original conformation. Those collapsed in the open channels are likely to remain distorted.

This proposal is fully substantiated by a constant pressure/constant enthalpy (N, P, H)^{24,25} molecular dynamics calculation on a model D3C clathrasil (with eight spherical guests to mimic the THF molecules). The calculated results are depicted in Fig. 3. The onset of the amorphization was found at $\sim 7 \text{ GPa}$, in reasonable agreement with experiment. The structure at 10 GPa, although heavily distorted, still shows remnants of crystallinity. When pressure is gradually lowered from 15 GPa, the original crystalline structure is fully recovered. In contrast, a similar molecular dynamics calculation on the empty clathrasil shows that the structure of the amorphous high-pressure phase is preserved after the removal of pressure.

The reversible amorphization effect in other materials may be rationalized in a similar manner. For example, the mineral berlinite is composed of interlinking AlO_4 and PO_4 tetrahedra in a α -quartz structure^{3,26}. Molecular dynamics calculations⁴ have shown that the helical AlO_4 and PO_4 tetrahedra are compressed and twisted as a response to the external stress, until the weaker AlO_4 tetrahedra distort so severely that the crystalline structure is no longer stable; the framework then distorts to give a higher density disordered phase at 20 GPa (ref. 3). The more rigid PO_4

units, however, remain intact and direct the transformed structure back into the crystalline form on release of the pressure. A similar explanation may be applied to other materials such as SnI_4 (refs 27, 28), LiKSO_4 (ref. 29) and Ca(OH)_2 (ref. 30), because they are composed of very open structures and possess strongly bonded tetrahedral or octahedral units. □

Received 2 February; accepted 9 May 1994.

- Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **310**, 393–395 (1984).
- Handa, Y. P., Tse, J. S., Klug, D. D. & Whalley, E. *J. Chem. Phys.* **94**, 623–627 (1991).
- Kruger, M. B. & Jeanloz, R. *Science* **249**, 647–649 (1990).
- Tse, J. S. & Klug, D. D. *Science* **255**, 1559–1561 (1992).
- Whalley, E., Mishima, O., Handa, Y. P. & Klug, D. D. *Ann. N.Y. Acad. Sci.* **484**, 81–92 (1986).
- Davidson, D. W. in *Water: A Comprehensive Treatise* Vol. 2 (ed. Franks, F.) 115–234 (Plenum, New York, 1973).
- van der Waals, J. H. & Platteeuw, J. C. *Adv. chem. Phys.* **2**, 1–57 (1959).
- Gies, H., Liebau, F. & Gerke, H. *Angew. Chem.* **94**, 214–216 (1984).
- Gies, H., Z. *Kristallogr.* **167**, 73–82 (1984).
- Hemley, R. J. et al. *Nature* **334**, 52–54 (1988).
- Gies, H. Z. *Kristallogr.* **175**, 93–104 (1986).
- Etchepore, J., Merian, M. & Kaplan, P. *J. chem. Phys.* **60**, 1873–1876 (1974).
- Bell, R. J. & Dean, P. *Discuss. Faraday Soc.* **50**, 55–60 (1970).
- Glinemann, J. et al. Z. *Kristallogr.* **198**, 177–212 (1992).
- Meade, C., Hemley, R. J. & Mao, H. K. *Phys. Rev. Lett.* **69**, 1387–1390 (1992).
- Klug, D. D., Mishima, O. & Whalley, E. *J. chem. Phys.* **86**, 5323–5328 (1987).
- Tse, J. S., Ripmeester, J. A. & Handa, Y. P. *J. Am. chem. Soc.* **115**, 281–284 (1993).
- Williams, Q. & Jeanloz, R. *Nature* **338**, 413–415 (1989).

19. Tse, J. S. & Klug, D. D. *Phys. Rev. Lett.* **67**, 3559–3562 (1991).
20. Tse, J. S. *J. chem. Phys.* **96**, 5482–5487 (1992).
21. Wolf, D., Okamoto, P. R., Lutsko, J. F. & Kluge, M. *J. Mater. Res.* **5**, 286–301 (1990).
22. Born, M. & Huang, K. *The Dynamical Theory of Crystal Lattices* 140–154 (Oxford Univ. Press, London, 1962).
23. Tse, J. S. & Klein, M. L. *Phys. Rev. Lett.* **58**, 1672–1674 (1987).
24. Rahman, A. & Parrinello, M. *J. appl. Phys.* **52**, 7182–7190 (1981).
25. Nosé, S. & Klein, M. *Molec. Phys.* **50**, 1055–1076 (1983).
26. Wyckoff, R. W. G. *Crystal Structures* Vol. 3 p. 31 (Interscience, New York, 1951).
27. Fujii, Y., Kowaka, M. & Onodera, A. *J. Phys. C* **18**, 789–797 (1985).
28. Sugai, S. *J. Phys. C* **18**, 799–808 (1985).
29. Melo, F. E. A., Lemos, V., Cerdiera, F. & Mendes, F. *Phys. Rev. B* **35**, 3633–3636 (1987).
30. Kruger, M. B., Williams, Q. & Jeanloz, R. *J. chem. Phys.* **91**, 5910–5915 (1989).
31. Klug, D. D. & Whalley, E. *Rev. Sci. Instrum.* **54**, 1205–1208 (1983).
32. Decker, D. L. *J. appl. Phys.* **42**, 3239–3244 (1971).
33. Cox, D. E., Hastings, J. B., Cardoso, L. P. & Finger, L. W. in *High Resolution Powder Diffraction* (ed. Catlow, C. R. A.) 1–20 (Trans Tech, Aedermannsdorf, Switzerland, 1986).
34. Fisher, J., Radeka, V. & Smith, G. C. *Nucl. Instrum. Meth.* **A252**, 239–245 (1986).
35. Mao, H. K. & Bell, P. M. *Carnegie Inst. Wash. Yb.* **79**, 409–411 (1980).
36. van Beest, B. W. H., Kramer, G. J. & van Santen, R. A. *Phys. Rev. Lett.* **64**, 1955–1958 (1990).

ACKNOWLEDGEMENTS. We acknowledge the help of D. E. Cox at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory and K. Brister at the Cornell High Energy Synchrotron Source for setting up the diffraction experiments. The NSLS is supported by the US Department of Energy.

Assembly of porphyrin building blocks into network structures with large channels

B. F. Abrahams, B. F. Hoskins, D. M. Michail & R. Robson

Inorganic Section, School of Chemistry, University of Melbourne, Parkville, Victoria 3052, Australia

CRYSTAL engineering—the deliberate design and construction of crystal structures from molecular components—promises to provide solid-state materials with specific and useful chemical, mechanical, electronic or optical properties¹. In most of the molecular crystals considered so far, van der Waals forces and hydrogen bonding govern the crystal packing^{2–7}. Zeolites, pillared clays and related microporous materials, which have been studied extensively because their porous structures convey useful catalytic activity^{8,9}, can now also be ‘engineered’ to some extent^{10,11}. We are exploring ways^{12–14} to construct channelled solids with very different chemical architectures and potentially different catalytic activity from those of zeolites. Here we show that porphyrin building blocks can be used to construct three-dimensional networks with the topology of the PtS structure, containing large channels. In our materials the channels are filled with solvent molecules, and crystalline order is lost on solvent removal. Nevertheless, the results show that it is possible to use simple molecular building blocks to engineer specific frameworks which, if they can be made robust, may offer new catalytic potential.

One approach to crystal engineering that we have been developing¹² is first to choose a geometrical/topological model one of a number of simple three-dimensional nets such as diamond (all centres tetrahedral), α -Po (all centres octahedral), PtS (equal numbers of square-planar and tetrahedral centres), rutile (octahedral and trigonal centres in 1:2 proportions) and so on, and then try to devise ways of chemically linking together molecular building blocks with a functionality and a stereochemistry appropriate to the chosen parent net. On a practical level, a working hypothesis we have adopted is that self-assembly to yield ordered infinite frameworks with the intended topology of the parent net may occur spontaneously when suitably functionalized and organized building blocks are brought into appropriate reaction. The success or failure of such framework-building attempts may be crucially dependent on the degree of organization built into the components to be united, but when this organization is sufficient the formation of some

pre-determined three-dimensional network may well become inevitable.

Frameworks based on the diamond structure might be expected to form if suitably functionalized tetrahedral building blocks could be built and appropriately connected. The linking together of 4,4',4'',4'''-tetracyanotetraphenylmethane units via tetrahedral Cu(I) centres provides an example of the deliberate construction of such a diamond-related network¹³. The PtS prototype net consists in essence of equal numbers of square-planar and tetrahedral centres, each one being connected to four of the other type. This net is of particular interest because it appears to be the simplest one to lend itself to the use of appropriately functionalized porphyrins and phthalocyanines as square-planar building blocks. To test as simply as possible the hypothesis that square-planar and tetrahedral units might spontaneously assemble themselves into a PtS-related net, in previous work we used $\text{Pt}(\text{CN})_4^{2-}$ as the square unit and Cu(I) ions as the tetrahedral unit; the resulting three-dimensional network did indeed have the intended PtS topology¹⁴. The present results are an extension of this work in which, again, Cu(I) is intended to provide the tetrahedral centre but now the square-planar building block is provided by the very much larger units 5,10,15,20-tetra(4-pyridyl)-21H,23H-porphine copper(II) (hereafter Cu(tpp)) or 5,10,15,20-tetrakis(4-cyanophenyl)-21H,23H-porphine copper(II) (hereafter Cu(tcp)).

Metalloporphyrin- and phthalocyanine-containing polymeric materials and molecular assemblies have aroused much interest and study^{15–20}. Until now, however, the only example of an infinite three-dimensional net constructed from 4-connecting porphyrin units was one in which the tendency of metalloporphyrins (lacking axial ligands) to π -stack dominated the arrangement and precluded the generation of large channels²¹.

When solutions containing Cu(I)(CH₃CN)₄BF₄ and either Cu(tcp) or Cu(tpp) in acetonitrile/nitrobenzene are allowed partially to evaporate under a stream of nitrogen, dark ruby-red crystals are formed. In the Cu(tcp) case the product was homogeneous, allowing analytical and infrared data to be obtained; but in the Cu(tpp) case, the well-formed crystals (characterized as described below by X-ray crystallography) were accompanied by unidentified dark amorphous material. The crystals of both products could be seen under the microscope to deteriorate on exposure to the air, quickly losing their single-crystal X-ray-diffracting behaviour. X-ray crystallographic analyses were therefore performed on crystals sealed in Lindemann glass tubes containing mother liquor.

In the Cu(tpp) case, despite the fact that the crystals were very weakly reflecting, the structural analysis proved that the intended PtS-related $\{[\text{Cu}(\text{II})(\text{tpp})\text{Cu}(\text{I})]_n\}^{++}$ framework had assembled. This framework comprised less than half the volume of the crystal, the remaining space being occupied by highly disordered solvent molecules and anions whose positions could not be defined. The arrangement of the framework atoms within the tetragonal unit cell is indicated in Fig. 1A, and a view of the extended framework structure revealing very large channels is shown in Fig. 1B: the highly disordered solvent and anions occupying these channels are not represented in the figures. An indication of the spaciousness of the structure is provided by the separation of ~ 14.2 Å between porphyrins facing each other across a channel (either from top to bottom as viewed in Fig. 1 or, as is equivalent, from front to back); taking a generous average van der Waals ‘half thickness’ of ~ 1.9 Å for the metalloporphyrin, this corresponds to a ‘free space’ gap of at least 10 Å.

The X-ray-diffraction study of the crystals obtained from the Cu(tcp)/CuBF₄ reaction reveals that here also an infinite PtS-related net has been formed; in this case, however, two such independent but interpenetrating $\{[\text{Cu}(\text{II})(\text{tcp})\text{Cu}(\text{I})]_n\}^{++}$ frameworks are present. The framework atoms within the unit cell are shown in Fig. 2A and a schematic representation of the extended interpenetrating pair of nets is shown in Fig. 2B. There are twelve Cu(tcp) units in Fig. 2A, six from each of two

independent frameworks. The individual frameworks have the same connectivity and almost the same geometry as the single $\{[\text{Cu}(\text{II})(\text{tcp})\text{Cu}(\text{I})]_n\}^{n+}$ framework represented in Fig. 1, which is viewed from almost the same angle. Understanding the doubled-up arrangement in Fig. 2 may be facilitated by comparison with the single framework in Fig. 1. The two independent networks are held, presumably by some attractive interaction, in close proximity throughout their three-dimensional entirety so that a system of large intersecting channels runs through the doubled-up framework. Again, the channels are occupied by highly disordered solvent and anions which could not be located in the structure analysis. The porphyrins appear in close pairs, one from each framework, in which the nearest atoms to each $\text{Cu}(\text{II})$ are two pyrrolic carbons of the neighbouring porphyrin at 3.43(6) Å. The porphyrins within a pair are displaced relative to each other so that the $\text{Cu}(\text{II}) \dots \text{Cu}(\text{II})$ separation is 4.86(2) Å, the $\text{Cu}(\text{II}) \dots \text{Cu}(\text{II})$ vector being inclined at $\sim 45^\circ$ to the normal to the porphyrin mean planes. The macrocycles are distorted from planarity so that the outer faces of a pair are somewhat concave; this is reflected in the fact that the $\text{Cu}(\text{I}) \dots \text{Cu}(\text{I})$ separation within each of the four equivalent pairs of $\text{Cu}(\text{I})$ s directly attached to any particular porphyrin pair is 5.62(2) Å, significantly greater than the $\text{Cu}(\text{II}) \dots \text{Cu}(\text{II})$ separation of 4.86(2) Å.

An indication of the spaciousness of this structure is provided by the fact that the distance between the vertical C-C 'edges' of the two $\text{Cu}(\text{tcp})$ units facing the reader in Fig. 2A is 24.39(7) Å;

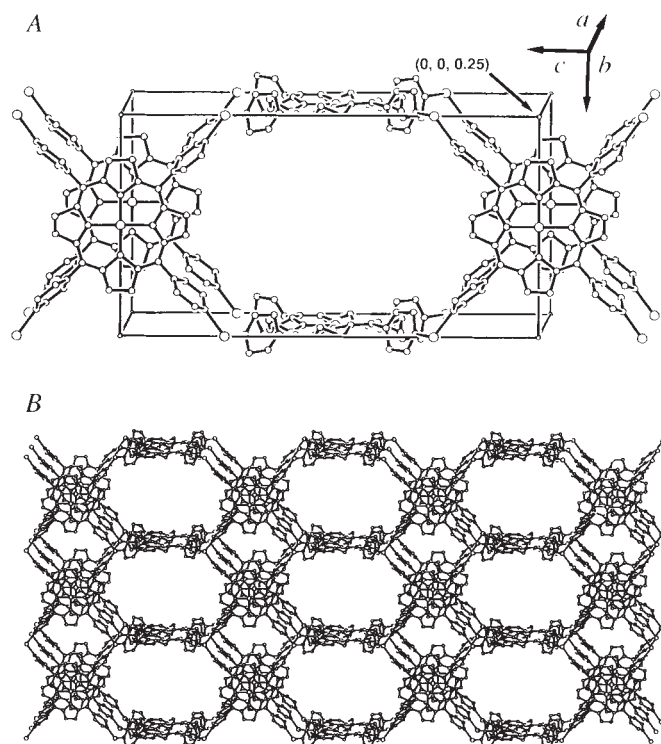


FIG. 1 A, The tetragonal unit cell and B, the extended structure of the $\{[\text{Cu}(\text{II})(\text{tcp})\text{Cu}(\text{I})]_n\}^{n+}$ framework. Larger circles represent copper atoms and smaller circles represent carbon and nitrogen atoms. The disordered solvent molecules, which fill the channels, are not shown here. METHODS. The structure was solved and refined (using the programs SHELXS-86 (ref. 25) and SHELXS-76 (ref. 26)) from data collected from a crystal of solvated $\text{Cu}(\text{tcp})\text{CuBF}_4$ using Mo $K\alpha$ radiation (0.71069 Å, graphite monochromator). Space group $P4_2c$; cell dimensions: $a = 14.202(6)$, $c = 26.644(9)$ Å. Maximum residual electron density, 0.48 electrons $\text{\AA}^{-3}$. Final $R = 0.112$ for 529 unique reflections ($I > 3\sigma(I)$ and $\theta > 7.5^\circ$) and 54 parameters; goodness of fit, 1.99. Further details of the data collection and structure solution are provided as Supplementary Information.

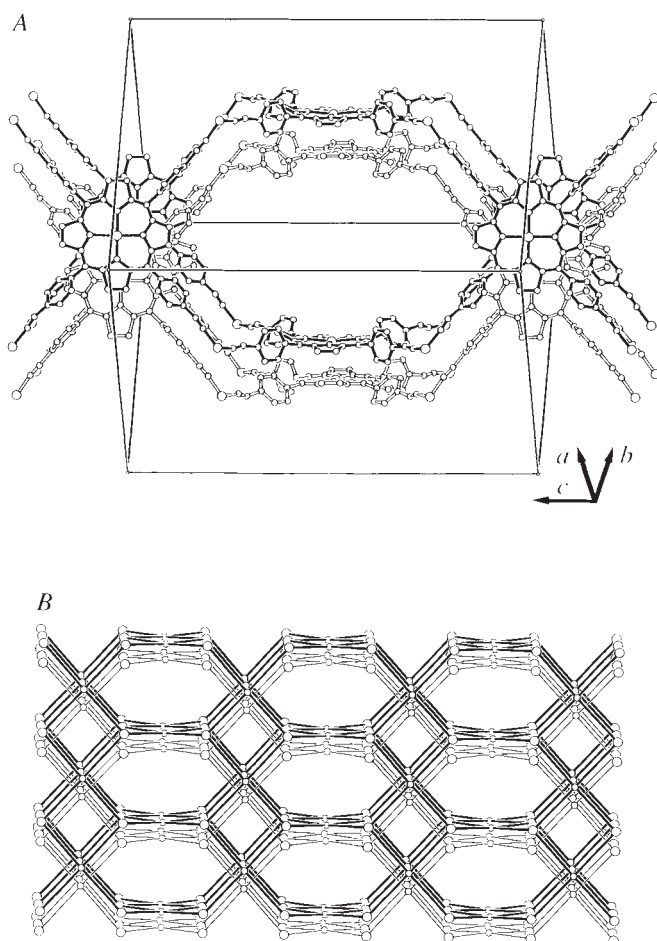


FIG. 2 A, The two independent, interpenetrating $\{[\text{Cu}(\text{II})(\text{tcp})\text{Cu}(\text{I})]_n\}^{n+}$ frameworks and the orthorhombic unit cell. Separate frameworks are indicated by heavy-line and open-line bonds. Large circles represent copper atoms and small circles represent carbon and nitrogen atoms. For clarity, each macrocyclic unit in the vertical plane is represented so that no structural features directly behind it are visible. As in Fig. 1, solvent molecules fill the channels. B, Representation of the two extended, independent, interpenetrating $\{[\text{Cu}(\text{II})(\text{tcp})\text{Cu}(\text{I})]_n\}^{n+}$ nets showing $\text{Cu}(\text{I}) \dots \text{Cu}(\text{II})$ connectivities. $\text{Cu}(\text{I})$ centres are represented by larger circles, $\text{Cu}(\text{II})$ centres by smaller circles. One framework is represented by heavy lines, the other by fine lines.

METHODS. The structure was solved and refined (using the programs SHELXS-86 (ref. 25) and SHELXS-76 (ref. 26)) from data collected from a crystal of solvated $\text{Cu}(\text{tcp})\text{CuBF}_4$ using Mo $K\alpha$ radiation (0.71069 Å, graphite monochromator). Space group $Cccm$; cell dimensions: $a = 23.69(1)$, $b = 27.26(2)$, $c = 32.76(1)$ Å. Maximum residual electron density, 0.65 electrons $\text{\AA}^{-3}$. Final $R = 0.142$ for 457 unique reflections ($I > 3\sigma(I)$ and $\theta > 7.5^\circ$) and 104 parameters; goodness of fit = 2.14. Further details of the data collection and structure solution are provided as Supplementary Information. The infrared spectrum (KBr disc) of the solid solvated $\text{Cu}(\text{tcp})\text{CuBF}_4$ indicated the presence of $\text{Cu}(\text{tcp})$ (1,000, 1,495, 1,595 cm^{-1}), BF_4^- (1,050 cm^{-1}) and nitrobenzene (705, 1,345, 1,520 cm^{-1}). The following analyses were conducted on a sample washed with nitrobenzene immediately after collection, and then immediately placed in an enclosed vessel containing an atmosphere saturated with nitrobenzene vapour; a fraction of the observed nitrobenzene present undoubtedly corresponds to liquid nitrobenzene external to the solid. Analysis: found; C, 59.6; H, 3.8; N, 11.8; Cu, 5.2; calculated for $\text{C}_{150}\text{H}_{109}\text{BCu}_2\text{F}_4\text{N}_{25}\text{O}_{34}$ (corresponding to $\text{Cu}(\text{tcp})\text{CuBF}_4 \cdot 17\text{C}_6\text{H}_5\text{NO}_2$); C, 59.7; H, 3.6; N, 11.6; Cu, 4.2%. Weight loss at 160 °C under vacuum, 69.3%. Calculated weight loss for removal of nitrobenzene from $\text{Cu}(\text{tcp})\text{CuBF}_4 \cdot 17\text{C}_6\text{H}_5\text{NO}_2$, 69.3%. Found for residue after drying at 160 °C under vacuum: C, 61.6; H, 3.0; N, 12.2; Cu, 13.8. Calculated for $\text{C}_{48}\text{H}_{24}\text{BCu}_2\text{F}_4\text{N}_8$ (corresponding to $\text{Cu}(\text{tcp})\text{CuBF}_4$); C, 62.2; H, 2.6; N, 12.1; Cu, 13.7%.

allowing for the van der Waals radii of the attached hydrogens, the 'free space' dimension of the channel in this direction (left to right in Fig. 2d) is close to 20 Å. The distance between a Cu(II) of one pair and the nearest Cu(II) of another pair across a channel is 15.64(2) Å.

We suggest that interpenetration of the sort observed for solvated Cu(tpc)CuBF₄ does not occur with solvated Cu(tpp)CuBF₄ because the required close approach of the Cu(I) centres from separate frameworks is precluded by the four bulky pyridyl units around each metal; by contrast it would be difficult to find four less sterically demanding donors than the nitrile ligands surrounding the Cu(I) in Cu(tpc)CuBF₄.

Paired metalloporphyrins have been studied as models for the 'special pair' of the photosynthetic system²²; paired frameworks related to that described here with a range of metallo-tpc units in place of Cu(tpc) could perhaps be made, and may furnish interesting photophysical and photochemical behaviour.

With regard to future studies, porphyrins and phthalocyanins are particularly alluring as building blocks for networks with potential applications as microporous heterogeneous catalysts because (1) their relative rigidity and large size, in the correct circumstances, could be used to generate correspondingly large channels and cavities, (2) they show a high degree of thermal stability and could therefore be put to use at elevated temperatures, (3) they readily incorporate a wide range of metal centres and (4) as discrete molecular species, they are known to catalyse and promote diverse processes²³. The porphyrin-containing networks described here are so delicate that they do not survive removal of nitrobenzene, so, at this early stage, materials of this general type fall far short of rivaling zeolites. We are presently exploring the possibility of putting to use the stability-enhancing chelate effect to provide more robust solids.

The exercise described here was conducted to demonstrate, in the simplest way we could visualize, what might be possible with porphyrin building blocks. The fact that the PtS network did assemble itself as intended is very encouraging with regard to future framework construction. On the basis of the results presented here, the deliberate construction of microporous heterogeneous catalysts (in which porphyrins and similar systems play the dual role of active site and framework component of sufficiently large size to produce large access channels) seems a realistic prospect. The basic approach is obviously amenable to wide variation; it seems feasible that appropriately designed building blocks may afford many nets other than PtS and diamond. With regard to the prospects for catalytic activity offered by framework solids in general, it is encouraging that a two-dimensional Cd(II)(4,4'-bipyridine)₂ sheet polymer containing corral-like square cavities surrounded by 4,4'-bipyridine 'fences'¹ has recently been shown to catalyse the cyanosilylation of aldehydes²⁴. □

17. Takemoto, K., Inaki, Y. & Ottenbrite, R. M. *Functional Monomers and Polymers* (Dekker, New York, 1987).
18. Collman, J. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4581–4585 (1986).
19. Gunter, M. J. et al. *Inorg. Chem.* **23**, 283–300 (1984).
20. Fleischer, E. B. & Shachter, A. M. *Inorg. Chem.* **30**, 3763–3769 (1991).
21. Abrahams, B. F., Hoskins, B. F. & Robson, R. J. *Am. chem. Soc.* **113**, 3606–3607 (1991).
22. Sessler, J. L., Johnson, A. R., Lin, T.-Y. & Creager, S. E. *J. Am. chem. Soc.* **110**, 3659–3661 (1988).
23. Walter, C. J., Anderson, H. L. & Sanders, J. K. M. *J. chem. Soc. chem. Commun.* 458–460 (1993).
24. Fujita, M., Kwon, Y. J., Washizu, S. & Ogura, K. *J. Am. chem. Soc.* **116**, 1151–1152 (1994).
25. Sheldrick, G. M. in *Crystallographic Computing Vol. 3* (eds Sheldrick, G. M., Kruger, C. & Goddard, R.) 175–189 (Oxford Univ. Press, 1985).
26. Sheldrick, G. M. *SHELXS-76, a Program for Crystal Structure Determination* (Univ. Cambridge, 1976).

SUPPLEMENTARY INFORMATION. Requests should be addressed to Mary Sheehan at the London editorial office of *Nature*.

ACKNOWLEDGEMENTS. We thank the Australian Research Council and the American Chemical Society Petroleum Research Fund for support.

Degradation of trifluoroacetate in oxic and anoxic sediments

Pieter T. Visscher, Charles W. Culbertson
& Ronald S. Oremland

US Geological Survey, Mailstop 465, 345 Middlefield Road,
Menlo Park, California 94025, USA

THE deleterious effect of chlorofluorocarbons on stratospheric ozone has led to international cooperation to end their use^{1–3}. The search for acceptable alternatives has focused on hydrofluorocarbons (HFCs) or hydrochlorofluorocarbons (HCFCs) which are attractive because they have relatively short atmospheric residence times⁴. HFCs and HCFCs are attacked by tropospheric hydroxyl radicals, leading to the formation of trifluoroacetate (TFA)⁵. Most of the atmospheric TFA is deposited at the Earth's surface⁶, where it is thought to be highly resistant to bacterial attack⁵. Therefore, use of HCFCs and HFCs may lead to accumulation of TFA in soils, where it could prove toxic or inhibitory to plants and soil microbial communities^{5,7}. Although little is known about the toxicity of TFA, monofluoroacetate, which occurs at low levels in some plants⁸ and which is susceptible to slow attack by aerobic soil microbes⁹, is known to be acutely toxic^{10–13}. Here we report that TFA can be rapidly degraded microbially under anoxic and oxic conditions. These results imply that significant microbial sinks exist in nature for the elimination of TFA from the environment. We also show that oxic degradation of TFA leads to the formation of fluoroform, a potential ozone-depleting compound with a much longer atmospheric lifetime than the parent compounds.

We chose to work initially with anoxic sediments because of their ability to degrade CFCs¹⁴ and other halogenated methyl compounds such as methyl bromide¹⁵. We incubated anoxic sediments from a San Francisco Bay saltmarsh¹⁶ and from a freshwater lake^{17,18} in the presence of 1.85 µM 2-¹⁴C-TFA (Fig. 1). Production of CH₄ (Fig. 1a) and ¹⁴CH₄ (Fig. 1b) in saltmarsh sediments only occurred when sulphate reduction was inhibited with molybdate¹⁹ or by elimination of sulphate from the slurry salts. No ¹⁴CH₄ was noted in the controls, which included autoclaved sediments containing molybdate (Fig. 1b), thereby reinforcing our conclusion that this was a microbial rather than chemical phenomenon. However, unmanipulated freshwater sediments (SO₄²⁻ < 1 mM) also readily produced ¹⁴CH₄ (Fig. 1b), thereby demonstrating that TFA degradation can occur in other types of sediments. The unlabelled carboxyl group of TFA was therefore cleaved to CO₂, as occurs in acetoclastic methanogenesis²⁰. Addition of 2-bromoethanesulphonic acid, an inhibitor of methanogenic bacteria¹⁹, blocked production of CH₄ and ¹⁴CH₄.

Received 29 December 1993; accepted 5 May 1994.

1. Bein, T. (ed.) *Supramolecular Architecture* 88–253 (ACS Symp. Ser. No. 499, Am. Chem. Soc., Washington DC, 1992).
2. Desiraju, G. R. *Crystal Engineering: The Design of Organic Solids* (Elsevier, Amsterdam, 1989).
3. Fagan, P. J. & Ward, M. D. *Scient. Am.* **267**, 28–34 (1992).
4. Simard, M., Su, D. & Wuest, J. D. *J. Am. chem. Soc.* **113**, 4696–4698 (1991).
5. Ermer, O. *J. Am. chem. Soc.* **110**, 3747–3754 (1988).
6. MacNicol, D. D., McKendrick, J. J. & Wilson, D. R. *Chem. Soc. Rev.* **7**, 65–87 (1978).
7. MacNicol, D. D. *Inclusion Compounds* Vol. 2, 1–45 (Academic, London, 1984).
8. Derouane, E. G., Lemos, F., Naccache, C. & Ribeiro, F. R. (eds) *Zeolite Microporous Solids: Synthesis, Structure and Reactivity* (NATO ASI Ser. C, Vol. 352, Kluwer Academic, Dordrecht, 1992).
9. Kerr, G. T. *Scient. Am.* **261**, 82–87 (1989).
10. Estermann, M., McKusker, L. B., Baerlocher, C., Merrouche, A. & Kessler, H. *Nature* **352**, 320–322 (1991).
11. Davis, M. E., Saldarriaga, C., Montes, C., Garces, J. & Crouter, C. *Nature* **331**, 698–699 (1988).
12. Bein, T. (ed.) *Supramolecular Architecture* 256–273 (ACS Symp. Ser. No. 499, Am. Chem. Soc., Washington DC, 1992).
13. Hoskins, B. F. & Robson, R. J. *Am. chem. Soc.* **112**, 1546–1554 (1990).
14. Gable, R. W., Hoskins, B. F. & Robson, R. J. *chem. Soc., chem. Commun.* 762–763 (1990).
15. Byrn, M. P. et al. *J. Amer. chem. Soc.* **115**, 9480–9497 (1993).
16. Wohrle, D. in *Phthalocyanines: Properties and Applications* (eds Leznoff, C. C. & Lever, A. B. P.) 55–132 (VCH, New York, 1989).

TABLE 1 Degradation of 2-¹⁴C-TFA in sediment slurries

Electron acceptor	TFA (μM)	Incubation on time (d)	¹⁴ CHF ₃ (μCi l ⁻¹)	¹⁴ CO ₂ (μCi l ⁻¹)	¹⁴ CH ₄ (μCi l ⁻¹)	TFA conversion (%)
O ₂	0.925	27	0 (0)	0	0	0 (0)
	0.463	27	0.60 (0.10)	0	0	2.4 (0.3)
	0.463	15	0.50 (0.35)	0	0	1.9 (1.5)
	0.185	15	2.55 (2.35)	0	0	25.5 (23.7)
NO ₃	0.925	15	0	0	0	0
	0.463	15	0	0	0	0
SO ₄ ²⁻	1.850	12	0	0	0	0
	0.925	18	0	0.05 (0.1)	0	0.1 (0.2)
None	0.463	18	0	3.05 (0.95)	0	12.2 (3.9)
	1.850	8	0	0	72.95 (3.25)	73.0 (3.3)
	0.925	15	0	0	34.95 (0.80)	69.9 (1.6)
	0.463	15	0	0	21.45 (0.70)	86.9 (2.8)

Experiments were carried out with 20 ml of slurry in 57-ml serum bottles (except oxic incubations, which were performed in 160-ml serum bottles). Oxic samples were pre-incubated by shaking with exposure to the atmosphere for 20 h in order to remove all potential reducing agents. Mean of three replicate slurries are given (figures in parenthesis indicate standard deviation). Autoclaved slurries, incubated anoxically with sulphate or under oxic conditions, did not produce any ¹⁴C-labelled gases, ruling out chemical processes.

Under these conditions, we did not observe the production of ¹⁴C-trifluoromethane, ¹⁴C-difluoromethane, ¹⁴C-methylfluoride or ¹⁴CO₂ in the headspaces of any of the slurries (not shown). We interpret these observations to mean that TFA is readily degraded under methanogenic conditions, but we were surprised that conditions that favoured sulphate-reduction did not result in formation of ¹⁴CO₂, which we would predict to occur for acetate²¹. We pursued the question of TFA mineralization by sulphate reduction in a subsequent experiment (see below). Because in our first experiment, the only end product was ¹⁴CH₄, it implied that TFA underwent a complete defluorination to the level of acetate before methanogenesis. To test this hypothesis, we examined the liquid phase, as well as the gas phase, of a larger volume of saltmarsh slurry incubated without sulphate. Results clearly demonstrated that a sequential order of defluorination occurred in these sediments (Fig. 2). Thus, the disappearance of 2-¹⁴C-TFA was respectively followed by the transient appearances of ¹⁴C-difluoroacetate (2-¹⁴C-DFA), ¹⁴C-mono-fluoroacetate (2-¹⁴C-MFA), 2-¹⁴C-acetate, and eventually the production of ¹⁴CH₄. By the time ¹⁴CH₄ production had ceased, after 7 days incubation, there no longer were detectable counts in the liquid phase. Mean conversion efficiencies of 2-¹⁴C-TFA to ¹⁴CH₄ was 80% for this experiment, and 66–77% in the previous experiments (see below). We attribute the slightly lower efficiencies to binding of acetate to sediment surfaces, a process which makes a portion of it unavailable to microbial degradation²², as the cause for the less than 100% recovery of 2-¹⁴C-TFA as ¹⁴CH₄. We did not detect any alteration of the 2-¹⁴C-TFA in autoclaved controls after prolonged (6 weeks) incubation (data not shown). Additionally, sediments inhibited by 2-bromoethanesulphonic acid had only a slight loss of TFA (15–20%) after 27 days incubation (data not shown). We conclude that this rapid defluorination is carried out by microorganisms, and the results with 2-bromoethanesulphonic acid suggest direct involvement of methanogens in the reductive defluorination of the TFA. The anaerobic microbial reductive dehalogenation of aromatic and aliphatic compounds has been well studied, but mostly centres on dechlorination and debromination reactions²³. Defluorination of fluoroaromatics occurs in methanogenic enrichments, but fluorine is removed before ring cleavage and acetate formation²⁴. This work is the first report of multiple reductive defluorinations of the methyl group of acetate.

Fluorinated compounds have been employed as microbial metabolic inhibitors^{13,18}. Monofluoroacetate, an intermediate of TFA degradation in our study, inhibited acetoclastic methanogenesis in lake sediments when applied at a concentration of ~20 μM (ref. 13). We observed an inhibitory effect of added unlabelled TFA on methanogenic activity at concentrations ≥1 μM. After 10 days incubation of sulphate-free saltmarsh slurries, mean methane levels (μmoles per 20 ml slurry; ±1 stan-

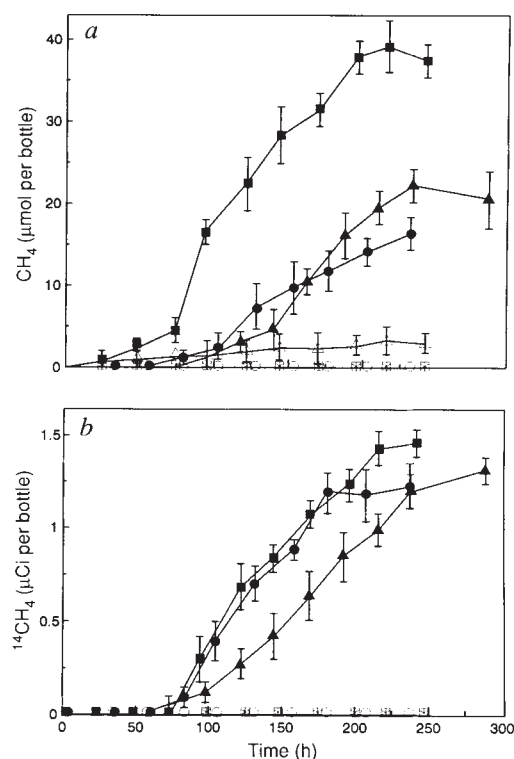


FIG. 1 Production of CH₄ (a) and ¹⁴CH₄ (b) in anoxic sediment slurries. Sediments were homogenized anaerobically with either artificial bay water (for saltmarsh sediments)²⁴ or with lake water (for fresh water)^{17,18} to yield final ratios of 1:3, sediment:water. Slurries (20 ml) were contained in 57-ml serum bottles sealed under O₂-free N₂ with black butyl crimp-seal stoppers. Samples received 2 μCi of 2-¹⁴C-TFA (Amersham Inc., Arlington Heights, Illinois); purity, 99.6%; specific activity, 54 mCi mmol⁻¹. Slurries were incubated in the dark at 23 °C with constant rotary shaking (200 r.p.m.). Gas phases were subsampled (250 μl) by syringe, and methane was analysed by flame-ionization gas chromatography³³, whereas ¹⁴CH₄ and ¹⁴CO₂ were analysed by gas chromatography in conjunction with proportional counting³⁴. Retention times were (min): CH₄ (0.8), CH₃F (2.5) and CHF₃ (3.5). At the end of the incubations, samples were injected with 2 ml of 6 M HCl and shaken overnight before determination of ¹⁴CO₂. Symbols: saltmarsh sediments incubated with 20 mM sulphate (*), without sulphate (■), with 20 mM sulphate plus 2.5 mM molybdate (●), without sulphate plus 5 mM bromoethanesulphonate (Δ), autoclaved with 20 mM sulphate and 2.5 mM molybdate (□), autoclaved with 20 mM sulphate (○); live freshwater sediments (▲). Symbols represent the mean of three individual slurries and error bars display ±1 standard deviation.

dard deviation; $n=3$) were; without additions (90 ± 15); plus $0.1 \mu\text{M}$ TFA (73 ± 7), plus $1 \mu\text{M}$ TFA (4.0 ± 2.1), and plus $10 \mu\text{M}$ TFA (2.4 ± 0.5).

Therefore, we hypothesized that our initial observation of no $^{14}\text{CO}_2$ production under conditions favouring sulphate-reduction may have been caused by stronger inhibitory effects of TFA (or of MFA) on sulphate-reducers. This led us to perform a series of experiments in which lower concentrations of TFA (0.185 – $0.925 \mu\text{M}$) were applied. The results from these incubations clearly show that sulphate-reducers can oxidize the methyl carbon of TFA to CO_2 , provided that the TFA concentration was lower than $0.925 \mu\text{M}$ (Table 1). Conversion efficiencies were highest at the lowest concentration ($0.463 \mu\text{M}$) of added TFA. We did not detect any degradation of TFA (that is, loss of TFA or formation of products) under conditions favouring nitrate respiration (Table 1), despite the fact that these slurries actively formed as much as 60% more unlabelled CO_2 than any of the other electron-acceptors tested (data not shown). However, with O_2 as electron acceptor, we detected a peak with a slightly shorter retention time than that of $^{14}\text{CO}_2$. We then used a 4 times longer column (7.3-m Porapak Q) which separated $^{14}\text{CHF}_3$ (retention time, 5.8 min) from $^{14}\text{CO}_2$ (retention time, 9.0 min), thereby confirming that the observed product was only $^{14}\text{CHF}_3$. We only observed production of $^{14}\text{CHF}_3$ when the concentration of TFA was $<0.463 \mu\text{M}$, with the highest conversion ratio at the lowest TFA concentration tested (Table 1). Collectively, our results demonstrate that trace levels of TFA can be degraded by reductive defluorination under anoxic conditions, and by decarboxylation under oxic conditions. Assuming that our lake and saltmarsh sediments are representative of processes occurring elsewhere in the biosphere, these results suggest that the microbial flora present in most soils and sediments have the capacity to attack and degrade TFA.

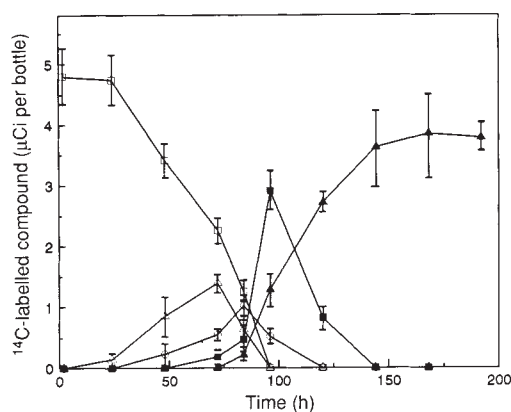


FIG. 2 Reductive defluorination of $2\text{-}^{14}\text{C}$ -TFA in saltmarsh sediment slurries incubated without sulphate. Preparation was the same as described for Fig. 1, except the volume was increased to 80 ml slurry in 158-ml serum bottles. Samples received $5 \mu\text{Ci}$ of $2\text{-}^{14}\text{C}$ -TFA. Slurry subsamples (1 ml) were removed by syringe and centrifuged in a micro-centrifuge (15,900g for 3 min). The supernatant was expressed through disposable $0.2 \mu\text{m}$ nylon filters and kept frozen until analysis (<3 weeks). Samples were analysed by high-performance liquid chromatography with an ultraviolet absorbance detector coupled with an inline β -emission radiation detector³⁵. Retention times determined with unlabelled standards were (min): TFA (4.5), difluoroacetate (6.0), monofluoroacetate (7.5) and acetate (9.5). There was a 1.2 min delay, determined by injection of $2\text{-}^{14}\text{C}$ -acetate and $2\text{-}^{14}\text{C}$ -TFA, between the mass detector and the radiation detector which allowed for identification and quantification with the use of a dual-pen chart recorder. Acetate concentration was $<30 \mu\text{M}$ throughout the experiment (not shown). Symbols: $2\text{-}^{14}\text{C}$ -TFA ($\square$), $2\text{-}^{14}\text{C}$ -difluoroacetate ($\triangle$), $2\text{-}^{14}\text{C}$ -monofluoroacetate ($\circ$), $2\text{-}^{14}\text{C}$ -acetate ($\blacksquare$) and $^{14}\text{CH}_4$ ($\blacktriangle$). Symbols represent the mean of three individual slurries and bars indicate ± 1 standard deviation.

Clearly, the concentration of TFA is a critical factor in this work. Our use of radiolabelled TFA enabled us to conduct experiments at concentrations that were at least four orders of magnitude lower than that employed previously by researchers using standard analytical chemical methods (F. G. Chumley, personal communication). The lowest concentration tested in our experiments was $0.185 \mu\text{M}$, based on the specific activity of the $2\text{-}^{14}\text{C}$ -TFA which we added. Although this is still ~ 1 – 2 orders of magnitude higher than that predicted for rainwater^{5,6}, we feel that degradation of TFA, especially if it is a co-metabolic feature^{25,26} is likely to occur unimpeded at even lower concentrations. Similarly, methanogenesis is inhibited by halogenated hydrocarbons^{27,28} like CCl_4 , although anaerobic degradation of CCl_4 occurs when the concentration $\leq 0.5 \mu\text{M}$ (refs 29, 30). Certainly our results with various electron acceptors indicated a trend of higher conversion efficiencies with decreasing TFA levels (Table 1). Although our present analytical techniques preclude conducting experiments at TFA concentrations $<0.1 \mu\text{M}$, we have no reason to question extrapolation of our results to anticipated natural concentrations. Still unanswered is the fate of TFA under conditions which favour other electron acceptors such as Fe^{3+} , NO_3^- and Mn^{4+} . However, our observation of TFA degradation under both oxic and anoxic conditions indicate that broadly-based microbial sinks exist for this compound in the terrestrial biosphere. Under anoxic conditions, TFA can be degraded to innocuous products like CH_4 and CO_2 . The fact that aerobic decarboxylation of TFA yields CHF_3 , a gas with a long atmospheric residence time, focuses attention on the fate of this product. That methane-oxidizing bacteria can degrade methylfluoride³¹ as well as HFC-143 (1,1,2-trifluoroethane) (ref. 32) suggest the possibility of a soil biological sink. $\square$

Received 14 December 1993; accepted 9 May 1994.

- Molina, M. & Rowland, F. S. *Nature* **249**, 810–812 (1974).
- Prather, M. J. & Watson, R. T. *Nature* **344**, 729–734 (1992).
- Global Ozone Research and Monitoring Project (Rep. No. 20, World Meteorological Organization, Geneva, 1989).
- Nimitz, J. S. & Scagg, S. R. *Envir. Sci. Technol.* **26**, 739–743 (1992).
- Franklin, J. *Chemosphere* **27**, 1565–1601 (1993).
- Rodriguez, J. M., Ko, M. K. W., Sze, N. D. & Heisey, C. W. *Proc. AFEAS Workshop on Atmospheric Wet and Dry Deposition of Carbonyl and Haloacetyl Halides 25–32* (AFEAS, Washington DC, 1992).
- Ingle, L. M. *Proc. West Virginia Acad. Sci.* **40**, 1–11 (1968).
- Oelrichs, P. B. & McEwan, T. *Nature* **190**, 808–809 (1961).
- Wong, D. H., Kirkpatrick, W. E., King, D. R. & Kinnear, J. E. *Soil Biol. Biochem.* **24**, 833–838 (1992).
- Kun, E. *Citric Acid Cycle, Control and Compartmentation* (ed. Lowenstein, J. M.) 279–339 (Dekker, New York, 1969).
- Bong, C. L., Cole, A. J. L., Walker, J. R. L. & Peters, J. A. *Soil Biol. Biochem.* **11**, 13–18 (1979).
- Notman, P. N. Z. *J. Entomol.* **12**, 67–71 (1989).
- Cappenberg, T. E. & Prins, R. A. *Antonie van Leeuwenhoek* **40**, 457–469 (1974).
- Lovley, D. R. & Woodward, J. C. *Envir. Sci. Technol.* **26**, 925–929 (1992).
- Oremland, R. S., Miller, L. G. & Strohmaier, F. E. *Envir. Sci. Technol.* **28**, 514–520 (1994).
- Oremland, R. S., Marsh, L. M. & Polcin, S. *Nature* **296**, 143–145 (1982).
- Oremland, R. S. & Culbertson, C. W. *Nature* **358**, 421–423 (1992).
- Miller, L. G., Coutlakis, M. D., Oremland, R. S. & Ward, B. B. *Appl. Envir. Microbiol.* **59**, 2457–2464 (1993).
- Oremland, R. S. & Capone, D. G. *Adv. microb. Ecol.* **10**, 285–383 (1988).
- Oremland, R. S. in *Biology of Anaerobic Microorganisms* (ed. Zehnder, A. J. B.) 641–706 (Wiley, New York, 1988).
- Oremland, R. S. & Polcin, S. *Appl. Envir. Microbiol.* **44**, 1270–1276 (1982).
- Christensen, D. & Blackburn, T. H. *Mar. Biol.* **7**, 113–119 (1982).
- Mohn, W. W. & Tiedje, J. M. *Microbiol. Rev.* **56**, 482–507 (1992).
- Londry, K. L. & Fedorak, P. M. *Arch. Microbiol.* **160**, 137–143 (1993).
- Oremland, R. S. & Zehr, J. P. *Appl. Envir. Microbiol.* **52**, 1031–1036 (1986).
- Taylor, B. F., Hearn, W. L. & Pincus, S. *Arch. Microbiol.* **122**, 301–306 (1979).
- Bauchop, T. J. *Bacteriol.* **94**, 171–175 (1967).
- Wood, J. M., Kennedy, F. S. & Wolfe, R. S. *Biochemistry* **7**, 1707–1713 (1968).
- Bouwer, E. J. & McCarty, P. L. *Appl. Envir. Microbiol.* **45**, 1295–1299 (1983).
- Bouwer, E. J. & McCarty, P. L. *Envir. Sci. Technol.* **45**, 1286–1294 (1983).
- Oremland, R. S. & Culbertson, C. W. *Appl. Envir. Microbiol.* **58**, 2983–2992 (1992).
- DeFlaun, M. F., Ensley, B. D. & Steffan, R. J. *Biotechnology* **10**, 1576–1578 (1992).
- Oremland, R. S. *Appl. Envir. Microbiol.* **42**, 122–129 (1981).
- Culbertson, C. W., Zehnder, A. J. B. & Oremland, R. S. *Appl. Envir. Microbiol.* **41**, 396–403 (1981).
- Culbertson, C. W., Strohmaier, F. E. & Oremland, R. S. *Origins Life Evol. Biosphere* **18**, 397–407 (1988).

ACKNOWLEDGEMENTS. We thank F. G. Chumley, P. Crill, B. Z. Fathepure and B. F. Taylor for their comments, and E. I. du Pont de Nemours & Co. for the gift of $2\text{-}^{14}\text{C}$ -TFA. This work was supported by the NASA Upper Atmosphere Research Program and the US Geological Survey.

Failure of climate regulation in a geophysiological model

James E. Lovelock* & Lee R. Kump†

* Coombe Mill, St Giles on the Heath, Launceston PL15 9RY, UK

† Earth System Science Center & Department of Geosciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

THERE has been much debate about how the Earth responds to changes in climate—specifically, how feedbacks involving the biota change with temperature. There is in particular an urgent need to understand the extent of coupling and feedback between plant growth, global temperature and enhanced atmospheric concentrations of greenhouse gases. Here we present a simple, but we hope qualitatively realistic, analysis of the effects of temperature change on the feedbacks induced by changes in surface distribution of marine algae and land plants. We assume that algae affect climate primarily through their emission of dimethyl sulphide^{1–8} (which may influence cloud albedo), and that land plants do so by fixation of atmospheric CO₂ (refs 9–12). When we consider how the planetary area occupied by these two ecosystems varies with temperature, we find that a simple model based on these ideas exhibits three feedback regimes. In glacial conditions, both marine and terrestrial ecosystems provide a negative feedback. As the temperature rises to present-day values, algae lose their strong climate influence, but terrestrial ecosystems continue to regulate the climate. But if global mean temperatures rise above about 20 °C, both terrestrial and marine ecosystems are in positive feedback, amplifying any further increase of temperature. As the latter conditions have existed in the past, we propose that other climate-regulating mechanisms must operate in this warm regime.

If living organisms participate in climate regulation in an active and responsive way, they do so most probably as part of a tightly coupled system, which includes the biota, the atmosphere, the ocean and the crustal rocks. Within such a system, the growth of organisms changes environmental conditions and environmental change feeds back on growth. The growth rate of organisms increases exponentially from 0 °C to at least 30 °C under laboratory conditions when nutrients and other needs are not lacking⁷. Under natural conditions, growth is limited by the availability of nutrients, light and temperature and the production rate, which is the product of the growth rate and the supply rate of some specific resource. These processes are illustrated in Fig. 1a, where curve I shows the variation of the growth or production rate of organisms with temperature. This curve is a gaussian fit to the observation³⁴ that growth is low below 5 °C and above 40 °C, and maximal near 22 °C. Curve A reflects the limitation of marine algal growth by the formation of the thermocline at temperatures above 8 °C which prevents the upwelling of nutrients from richer waters below. Curve P shows how the growth rate of land plants might similarly be limited by water stress at temperatures near 18 °C.

We have modelled the coupling between growth and climate using equations similar to those of the Daisyworld model^{13,14}. The model parameter chosen to link the effect of temperature on growth, and the effect of growth on temperature, is the planetary area covered by land-plant and algal ecosystems (Fig. 1b). Note the steepness of the rise of area cover when the temperature reaches 6.5 °C for algae and 9 °C for land plants, redolent of the springtime bloom at sea and the greening of the land. In the model, the areal cover by plants and algae is directly linked to the albedo and to the size of the CO₂ sink. At temperatures for which the areal cover rises steeply with temperature, there will be strong negative feedback on temperature—extensive cover lowers the temperature. As the temperature for maximum areal cover is approached, the curve flattens, negative feedback lessens, and beyond the maximum it becomes positive.

What justification is there for the curves in Fig. 1? Riley and Chester¹⁵ have remarked that “it is a curious, and so far unexplained, fact that most species of phytoplankton grow best in culture at temperatures 5–10 °C higher than those at which optimum growth occurs in their natural habitat”. Satellite observations³⁵ have confirmed this, and have shown that lush algal growth is limited largely to ocean regions in which the surface temperature¹⁶ is near or below 8 °C; these appear also to be the regions of maximum cloud cover⁵. Algal ecosystems of the oceans are mainly found in cool regions, not because ocean organisms differ in their temperature response from other plants, but because the ocean surface stratifies as a thermocline when the heat flux increases¹⁷. Marine organisms in this surface layer are isolated from nutrients, which remain in the cooler waters below; thus, life is sparse in the tropical and warm temperate oceans.

Local studies^{18,19} have so far failed to find convincing evidence for an effect of algal growth on climate. But if the strong correlation between temperature, algal density and cloud albedo seen by satellite⁵ is no mere coincidence, then the present sparsity of dense algal cover, and of clouds such as marine stratus, could be a consequence of the warmer climate of the interglacial period, with temperatures well above the optimum for algae, set by the thermocline. In addition, evidence from Antarctic ice cores^{20,21} for the past two glacial cycles suggest a link between climate and atmospheric composition: during glacial times, the abundances of methane and of carbon dioxide are lower and that of non-sea-salt sulphate aerosol and of methane sulphonate (produced by oxidation of dimethyl sulphide) are higher^{19–22}. The timing and magnitude of the climate and chemistry record of the ice cores is independently confirmed by evidence from sediment cores in the oceans²³; these also suggest that at low latitudes the areal coverage of algae was greater in the glacial periods²⁴. This correlation is consistent with our model of self-regulation. We note, however, that ice-core data from Greenland show no significant change of methane sulphonate deposition between the glacial and the interglacial periods²⁵. This is not necessarily contradictory to our argument as only 30% of Greenland's coastline directly fronts the ocean, whereas for Antarctica the proportion is closer to 100%; moreover, in the last glacia-

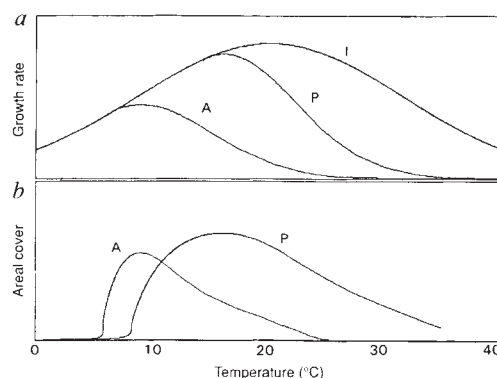


FIG. 1 Growth rate (specifically production rate) of, and areal cover by, land plants and ocean algae as a function of temperature. a, Variation of growth rate with temperature. Curve I is for ideal natural conditions where nutrients and water are not lacking, curve A is for ocean algae where the onset of the thermocline suppresses growth above 8 °C and curve P is for land plants where water stress limits growth above 18 °C. b, Variation with temperature of the planetary area covered by plants or algae; the area values were plotted directly from a numerical model that used the growth curves of a as input data. For both plants and algae the areal cover is shown as a fraction of the total available area and all effects are scaled relative to the maximum effect.

tion, pack ice extended in the Atlantic in winter as far south as the Canaries (30° N)²⁶, making Greenland even more distant from marine sources of sulphur gases and aerosols. If the Antarctic observations are representative, they suggest—like our postulated model of a self-regulating Earth—positive, not negative, feedback on climate change. The ubiquity of the thermocline today limits the area of ocean available to marine algae to sub-polar waters and to smaller areas where nutrients are abundant as a result of upwelling or riverine inputs. In all of these areas, local temperature regulation may take place, but their combined extent is now insufficient for planetary-scale regulation.

On the land surface (during the interglacial) there is less latitudinal limitation to growth than in the ocean, and the area (with average temperatures below 18 °C) available for plant growth may be sufficient to maintain climate regulation. In glacial periods, climates suitable for plant-mediated chemical weathering would have covered much of the available planetary surface. In the interglacial periods, climates are suitable in the temperate zones but in land regions between 30° N and 30° S desert and arid regions are prevalent. Where vegetation is abundant, as in the humid tropical forests, it may be a consequence of the forest ecosystem's ability to self-regulate and sustain locally lower temperatures. Humid tropical forests may not in themselves much affect the local rate of chemical weathering, but the effect of their presence on regional and global climate and rainfall may do. The global climate is therefore to some extent dependent on the area of land occupied by these forests. The fact that the clearance of tropical forests is apparently irreversible (at least on the timescale of human observation)²⁷ suggests that they are unsuited to the present interglacial period, and persist only by their capacity, when intact, to regulate their local physical environment. This proposal is supported by climate model simulations of the effects of deforestation^{28,29}. The rapidity of atmospheric mixing ensures that CO₂ changes are global, but CO₂ abundance is determined by local and regional sources and sinks.

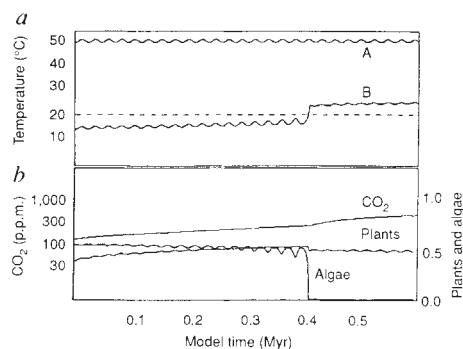


FIG. 2 A zero-dimensional flat Earth model where temperature regulation is coupled to the growth of plant and algal ecosystems. In the model, plant growth amplifies the rate of CO₂ removal from the atmosphere by chemical weathering and is optimal at 20 °C. Algal growth cools by cloud formation and is optimal at 10 °C. The initial conditions of carbon dioxide input and solar luminosity are those for the present-day Earth. Luminosity is that of the Sun, but modulated sinusoidally with a period of 20 kyr and amplitude of 1%, to show the presence of positive or negative feedback. *a*, Temperature change for a barren (curve A) and a live (curve B) Earth. *b*, Abundance of algae, plants and atmospheric carbon dioxide. Carbon dioxide input increases linearly, starting at the present geological input and rising three-fold. The amplitude of the temperature fluctuations for the barren state (A) can be used to diagnose the sign of the feedback. Note the onset of instability (larger amplitude fluctuations of temperature and algal abundance) leading to a transition to stable higher temperatures, when regulation is through the presence of plants alone. The area cover (curve B) for plants and algae is scaled to the maximum for land and ocean.

In the long time frame of the biologically assisted removal of CO₂ by rock weathering, any significant temperature increase above the optimum for plant growth is adverse and could lead to more desert and to reduced removal of CO₂. CO₂ removal by rock weathering is a slow process, with a time constant greater than 10⁵ yr (ref. 11)—fast enough to account for the low level of CO₂ at the end of the last glacial period, but too slow to explain the sudden rise of CO₂ into the interglacial. The positive feedback on CO₂ with rising temperature must have other causes: a likely candidate is the extension of the thermocline which decreases the area of ocean available to those algae that remove CO₂³⁰. The pumping down of CO₂ by ocean algal systems has a much shorter time constant than weathering, ~100–1,000 yr (ref. 36). CO₂ would rise as the algal pumping system declined, and as the anaerobic fermentation of vegetation on low-lying land flooded by sea water released CO₂ and CH₄ to the atmosphere. Rising sea levels may also have increased atmospheric CO₂ by similar processes occurring on coral reefs³¹. Together, these effects, and the atmospheric oxidation of the CH₄, would add a pulse of CO₂ to the atmosphere that would take thousands of years to remove, similar to the time it will take to remove the surge of CO₂ from fossil-fuel combustion³².

Figure 2 illustrates calculations of climate regulation based on a simple numerical model using the relationships depicted in Fig. 1. In this model the planetary albedo is coupled with the area of ocean occupied by algae, and the sink for CO₂ on the land surface is coupled with the land area occupied by plant growth. We have considered the effect of a progressive increase of the CO₂ source, going from the current volcanic input of 6×10^{12} mol yr⁻¹ to three times as great over a period of 500 kyr. Increasing CO₂, not solar luminosity, was chosen as the forcing function both because it illustrates better the failure of regulation and also because of its relevance today; the response of the model to increasing luminosity is similar except that the change in CO₂ abundance is less.

The stability of the model is measured by the extent to which cyclical variations of the input solar energy are damped; such variations are included in the model and show up as the small, superimposed oscillations. In the low-temperature regime (about 200 p.p.m. CO₂), both plants and algae are involved in regulation, and the system is stable. With increasing carbon dioxide, and consequently warmth, the area available for algal growth diminishes and the system becomes less stable. At about 400 p.p.m. CO₂ there is a sudden transition to a high-temperature regime, in which the algal ecosystem collapses and climate is regulated by land plants only, with the CO₂ concentration rising to 700 p.p.m. Before the algal system collapses, climate is regulated at temperatures higher than the optimum for plant growth (18 °C). This is because the planetary temperature is the average for the whole planet, and includes areas covered by cooler plants and hotter bare ground. As a caricature of the Earth, the model transition corresponds to the change from glacial to permanent interglacial. It suggests that the regulation system is more robust in the glacial periods, and that increasing CO₂ may stabilize a permanent warm state, hotter than today. Self-regulation in this warm state would depend critically on the extent to which the remaining organisms were able to accelerate chemical weathering^{12,33}.

The geophysiological view of climate change presented here ignores important geophysical feedbacks, such as the ice albedo effect and changes in ocean circulation. Simple world models tend to predict imminent disaster: the first simple ozone-depletion models predicted five times the depletion observed at present. The picture presented here of a 'malignant' greenhouse effect may be a similar exaggeration, but it serves to warn of the dangers of adding greenhouse gases to the atmosphere and destroying natural ecosystems at a time when the postulated geophysiological system itself may be dysfunctional, and when the consequences of these actions may be amplified by positive feedback. A more realistic model of the present Earth would

include zones of dense algal growth in the sub-polar oceans, and abundant plant growth in the humid tropical and the boreal forests. But most of the ocean and much of the land surface are no longer covered by these self-regulating ecosystems. In addition to the effects of higher temperatures on areal cover, nearly half of the land surface is now farmed. A knowledge of the area of ocean and land surface occupied by natural ecosystems in a state of healthy negative feedback is needed to estimate the extent of global climate self-regulation and the consequences for predictions of climate change. □

Received 4 August 1993; accepted 18 May 1994.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Coakley, J. A., Bernstein, R. L. & Durkee, P. A. *Science* **237**, 1020–1022 (1987).
3. Ayers, G. P. & Gras, J. L. *Nature* **353**, 834–835 (1991).
4. Andreae, M. O. in *The Role of Air-Sea Exchange in Geochemical Cycling* (ed. Buat-Menard, P.) 331–362 (Reidel, Dordrecht, 1986).
5. Falkowski, P. G. et al. *Science* **256**, 1311–1313 (1992).
6. Shaw, G. *Clim. Change* **5**, 297–303 (1983).
7. Lawrence, M. G. *J. geophys. Res.* **98**, 20663–20673 (1993).
8. Holligan, P. M. in *Primary Productivity and Biogeochemical Cycles in the Sea*, (eds Falkowski, P. G. & Woodhead, A. D.) 487–501 (Plenum, New York, 1992).
9. Lovelock, J. E. & Whitfield, M. *Nature* **296**, 561–563 (1983).
10. Lovelock, J. E. & Watson, A. J. *Planet. Space Sci.* **30**, 795–802 (1982).
11. Berner, R. A. *Geochim. cosmochim. Acta* **56**, 3225–3231 (1992).
12. Schwartzmann, D. W. & Volk, T. *Nature* **340**, 457–460 (1989).

13. Lovelock, J. E. *Rev. Geophys.* **27**, 215–222 (1989).
14. Watson, A. J. & Lovelock, J. E. *Tellus* **35B**, 284–289 (1983).
15. Riley, J. P. & Chester R. in *Introduction to Marine Chemistry* 42 (Academic, London, 1971).
16. Berger, W. H., Fischer, K., Lai, C. & Wu, G. Report No. 87–30 (Scripps Inst. Oceanogr., San Diego, 1987).
17. Neumann, G. & Pierson, W. J. Jr *Principles of Physical Oceanography* (Prentice-Hall, New Jersey, 1966).
18. Sievering, H. et al. *Nature* **360**, 571–573 (1992).
19. Charlson, R. J., Langer, J. & Rhode, H. *Nature* **348**, 22–25 (1990).
20. Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **329**, 408–414 (1987).
21. Genthon, C. *Nature* **329**, 414–418 (1987).
22. Legrand, M. et al. *Nature* **350**, 144–146 (1992).
23. Shackleton, N. J. & Opdyke, N. D. *Quater. Res.* **3**, 33–55 (1973).
24. Berger, W. H., Smetacek, V. S. & Wefer, G. *Productivity of the Ocean: Present and Past* (Wiley Interscience, New York, 1989).
25. Whung, P. Y., Salzman, E. S. & Davis, M. D. *EOS* **74**, 84 (1993).
26. CLIMAP Project Members *Science* **191**, 1131–1137 (1976).
27. Whitmore, T. C. *An Introduction to Tropical Rain Forests* (Oxford Univ. Press, 1992).
28. Dickinson, R. E. (ed.) *The Geophysics of Amazonia* (Wiley, New York, 1987).
29. Lean, J. & Warrilow, D. A. *Nature* **342**, 411–413 (1989).
30. Watson, A. J., Robinson, C., Robertson, J. E., Williams, P. J. leB. & Fasham, M. J. R. *Nature* **350**, 50–53 (1991).
31. Opdyke, B. N. & Walker, J. C. G. *Geology* **20**, 733–736 (1992).
32. Walker, J. C. G. & Kasting, J. F. *Glob. Planet. Change* **5**, 151–189 (1992).
33. Berner, R. A. *Am. J. Sci.* **294**, 56–91 (1994).
34. Bannister, P. *Introduction to Physiological Plant Ecology* (Wiley, New York, 1976).
35. Lewis, M. R. *New Scientist* 7 October (1989).
36. Sundquist, E. T. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Recent* (eds Sundquist, E. T. & Broecker, W. S.) 5–60 (American Geophysical Union, Washington DC, 1985).

ACKNOWLEDGEMENTS. This work was supported by Gaia Charity using funds donated by the World City Corporation. L.R.K. thanks Shell Research Ltd and the US NSF for support.

Response of the climate system to atmospheric aerosols and greenhouse gases

K. E. Taylor & J. E. Penner*

Global Climate Research Division, Lawrence Livermore National Laboratory, PO Box 808, Livermore, California 94551–990, USA

RECENTLY, Kiehl and Briegleb¹ evaluated the radiative forcing associated with the capacity of atmospheric sulphate aerosols to reflect solar radiation back into space, and compared this with the forcing associated with atmospheric greenhouse gases. They found that the (negative) climate forcing by the aerosols has strong regional character, with the greatest forcing over Northern Hemisphere land surfaces, whereas the (positive) forcing by greenhouse gases is distributed almost equally between the hemispheres and varies mainly as a function of latitude. Here we present simulations of the response of the climate system to these two types of forcing. We find that the global response to aerosol forcing is regionally heterogeneous, with a distribution that is different from the forcing pattern. The simulations also imply that, for equal magnitudes of forcing, the temperature response is markedly greater for carbon dioxide than for aerosol forcing. We conclude that to predict the global mean climate response to global mean forcing, it is necessary to separate out the different components of the forcing to which the climate system is sensitive.

Anthropogenic emissions of sulphur dioxide from industrial activity, fossil-fuel combustion and biomass burning are now known to be large enough (relative to natural sources) to produce significant increases in sulphate aerosol abundances. The resulting increased particle concentration in the atmosphere may decrease the globally-averaged solar radiation absorbed by the Earth–atmosphere system by as much as 1 W m^{-2} (ref. 2). This direct effect of aerosol particles may be supplemented by an indirect effect, in which aerosol particles act as cloud condensa-

tion nuclei (CCN). Increases in CCN would create clouds with more (but smaller) droplets, increasing further the reflection of solar radiation^{3,4}, but the magnitude of this indirect effect is unknown, and it will not be considered further here.

The climate effects resulting from a combination of sulphate aerosol and greenhouse-gas forcings have not been previously examined through model simulations, although analysis of observed temperature change suggests that the aerosol forcing is qualitatively consistent with the view that anthropogenic sulphate aerosols cause a substantial cooling in regions that are associated with the main anthropogenic emissions and sulphate burden^{5,6}. Here we report the results of general circulation model (GCM) calculations to assess the anticipated patterns of temperature response to CO₂ and the direct effects of a sulphate-aerosol increase.

We used the Lawrence Livermore National Laboratory tropospheric chemistry model (GRANTOUR⁷) in conjunction with the Livermore version of the NCAR CCM1 model⁸, a general circulation model (GCM) which was coupled to a 50-m mixed-layer ocean model with sea ice and specified meridional oceanic heat flux⁹. The model produces a reasonable, but somewhat cold present-day climate (Table 1) due primarily to a dry bias in CCM1 in the upper troposphere. The annual average shortwave and longwave cloud radiative forcing is in excellent agreement with the Earth Radiation Budget Experiment (ERBE) data, but the seasonal cycle of the cloud forcing is too weak. The response of the cloud radiative forcing to uniform increases in sea surface temperature (from 2 °C below the climatological average to 2 °C above) is similar to that of other climate models, and consequently the climate sensitivity ($0.76 \text{ °C W}^{-1} \text{ m}^2$) calculated from these runs for perpetual July conditions is within the range estimated by most climate models¹⁰.

The chemical model used here treats the gas-phase reaction of dimethyl sulphide and SO₂ with the OH radical, and assumed the latitudinally- and seasonally-varying OH field that was used in a previous study of the nitrogen cycle¹¹. The aqueous conversion rate of SO₂ to sulphate varies with season, proportional to the square of the local concentration of OH. This represents a simplified attempt to simulate the limiting reaction of H₂O₂ with SO₂ in clouds. In simulations using fixed sea surface temperatures, our model provides a reasonable account of the observed variations in both sulphate concentrations and deposition in precipitation in the eastern United States¹². The

* To whom correspondence should be addressed.

TABLE 1 Radiative forcing and climate statistics

Case	ΔF (W m^{-2})	T_s ($^{\circ}\text{C}$)	ΔT_s ($^{\circ}\text{C}$)	ΔP (mm d^{-1})	P (mm d^{-1})	C (%)	ΔC (%)	SI (%)	ΔSI (%)
Northern Hemisphere									
Pre-industrial		12.5		3.40		56.6		4.87	
Present-day CO_2	1.26	14.5	1.9	3.48	0.09	55.0	-1.7	4.13	-0.74
Present-day sulphate	-1.60	11.3	-1.2	3.36	-0.04	56.9	0.3	5.54	0.67
Combined CO_2 and sulphate	-0.34	13.0	0.5	3.43	0.03	55.8	-0.9	4.85	-0.02
Observed climate statistics		14.9*		2.6†		58.9‡		4.4§	
Southern Hemisphere									
Pre-industrial		12.5		3.54		62.4		6.64	
Present-day CO_2	1.25	14.8	2.3	3.61	0.08	61.1	-1.3	4.39	-2.26
Present-day sulphate	-0.30	11.7	-0.8	3.48	-0.06	63.1	0.7	7.24	0.59
Combined CO_2 and sulphate	0.95	13.6	1.1	3.56	0.02	62.1	-0.3	5.40	-1.24
Observed climate statistics		13.5‡		2.7†		65.6‡		4.5§	
Global average									
Pre-industrial		12.5		3.47		59.5		5.76	
Present-day CO_2	1.26	14.6	2.1	3.55	0.08	58.0	-1.5	4.26	-1.50
Present-day sulphate	-0.95	11.5	-1.0	3.42	-0.05	60.0	0.5	6.39	0.63
Combined CO_2 and sulphate	0.31	13.3	0.8	3.49	0.02	58.9	-0.6	5.13	-0.63
Observed climate statistics		14.2‡		2.7†		62.2‡		4.5§	

Simulated and observed annual hemispheric or global mean values of radiative forcing ΔF and climate statistics (surface air temperature T_s , precipitation P , cloud cover C and sea ice coverage SI). Also shown are the differences between the three cases with sulphate aerosol and/or CO_2 forcing, and the pre-industrial case.

* Ref. 21; † ref. 22; ‡ ref. 23; § ref. 24; ref. 25.

model also compares reasonably well with data from remote locations.

In the simulations, the anthropogenic sulphur inventories developed by Spiro *et al.*¹³ (for SO_2 from industrial, fossil fuel and biomass burning) were used, except over North America where the inventory developed by Benkowitz¹⁴ (for fossil fuel and industrial SO_2 emissions) was used. This anthropogenic inventory totals 78 Tg-S yr^{-1} , and represents annual average emissions for ~ 1980 .

In order to assess the anticipated patterns of temperature response to anthropogenic emissions of sulphur dioxide and carbon dioxide, we performed four calculations, each of 20 (model) years duration: (1) a pre-industrial case with no anthropogenic aerosol and 275 p.p.m. CO_2 ; (2) no anthropogenic aerosol and near-present-day atmospheric CO_2 concentration of 345 p.p.m. ; (3) present-day (1980) anthropogenic sulphur emissions and 275 p.p.m. CO_2 ; and (4) present-day anthropogenic sulphur emissions and 345 p.p.m. CO_2 . Table 1 shows the annual average radiative forcings and calculated climate statistics from the last 10 years of each simulation, along with the observed climatological means.

Our calculated forcing by anthropogenic sulphate aerosols in the combined climate-chemistry-ocean mixed-layer model is $\sim -0.9 \text{ W m}^{-2}$, a factor of ~ 3 larger than the recent calculation by Kiehl and Briegleb¹. Three factors contribute to our higher calculated forcing. First, whereas predicted sulphate concentrations are in reasonable agreement with data in most regions of the coupled mixed-layer ocean model, the predicted sulphate concentrations over Europe increased substantially over those predicted in the model with fixed sea surface temperatures¹²; they are larger than observations by a factor of ~ 2 because of changes in precipitation patterns that occur when the model is coupled to the mixed-layer ocean model. However, the region between 0° to 45° E and 30° to 60° N (which encompasses the area of overprediction) contributes only -0.1 W m^{-2} on a global average basis to the total forcing. Second, our simulated sulphate concentrations vary with season more strongly than the fields used by Kiehl and Briegleb¹, reflecting the strong observed seasonal variation that results from the aqueous and gas-phase chemistry that converts SO_2 emissions to sulphate aerosol. We estimate that the enhanced Northern Hemisphere summer concentrations of sulphate may increase our estimated forcing by $\sim -0.1 \text{ W m}^{-2}$. Third, our solar radiation calculation accounts for radiation transfer in seven wavelength bands varying from

$0.3 \mu\text{m}$ to $1.5 \mu\text{m}$. The scattering coefficient for wavelengths $>0.55 \mu\text{m}$ was decreased in proportion to the dry sulphate aerosol optical properties given by Kiehl and Briegleb¹. But whereas Kiehl and Briegleb varied the scattering coefficient of the aerosol depending on the local relative humidity, we assumed a constant scattering coefficient of $8.5 \text{ m}^2 \text{ g}^{-1}$ at $0.55 \mu\text{m}$, consistent with recommendations¹⁵ for the scattering properties of sulphate aerosol at an average relative humidity of 80%. Our calculated forcing decreases to -0.6 W m^{-2} if a relative-humidity-dependent scattering coefficient¹ is used. We believe that the higher average forcing calculated with the constant scattering coefficient may not be unrealistic for two reasons. First, the relative-humidity dependence of the scattering coefficient has not been unambiguously determined and is highly uncertain². Second, higher scattering effects by anthropogenic aerosols are anticipated because other anthropogenic aerosol components would be associated with the sulphate aerosol that is specifically modelled here¹³ and because the effects of aerosols on clouds have not been included.

Figure 1a and b shows the calculated annual average forcing and temperature response to the imposed increase in CO_2 . Note that the calculated forcing increases with increasing temperature (roughly according to the Stefan-Boltzman law). The temperature response, in contrast, is strongest at the poles. This feature, common to most climate models, is partly due to negative lapse rate feedbacks in the tropics, which reduce sensitivity there, and to the decrease in sea ice at polar latitudes, which contributes to a positive feedback in that region¹⁶.

Figure 1c and d shows the annual average forcing and temperature response for anthropogenic sulphate aerosols. The over-estimate of predicted sulphate concentrations in Europe noted earlier is apparent in the forcing patterns over Europe, which are considerably stronger than those calculated previously^{1,15}.

The temperature response to the sulphate forcing (a cooling) is stronger in the polar regions, again due, in part, to the increases in sea ice associated with the ocean model. The pattern of the model's predicted response for this case is similar to that presented by Roeckner¹⁷, in that the greatest cooling is found over broad regions of the Northern Hemisphere continents. The decrease in temperatures near the Antarctic sea-ice margins is not associated with a direct local forcing in that area but is caused by the general global cooling. The maximum cooling in the Northern Hemisphere does not coincide with the strongest

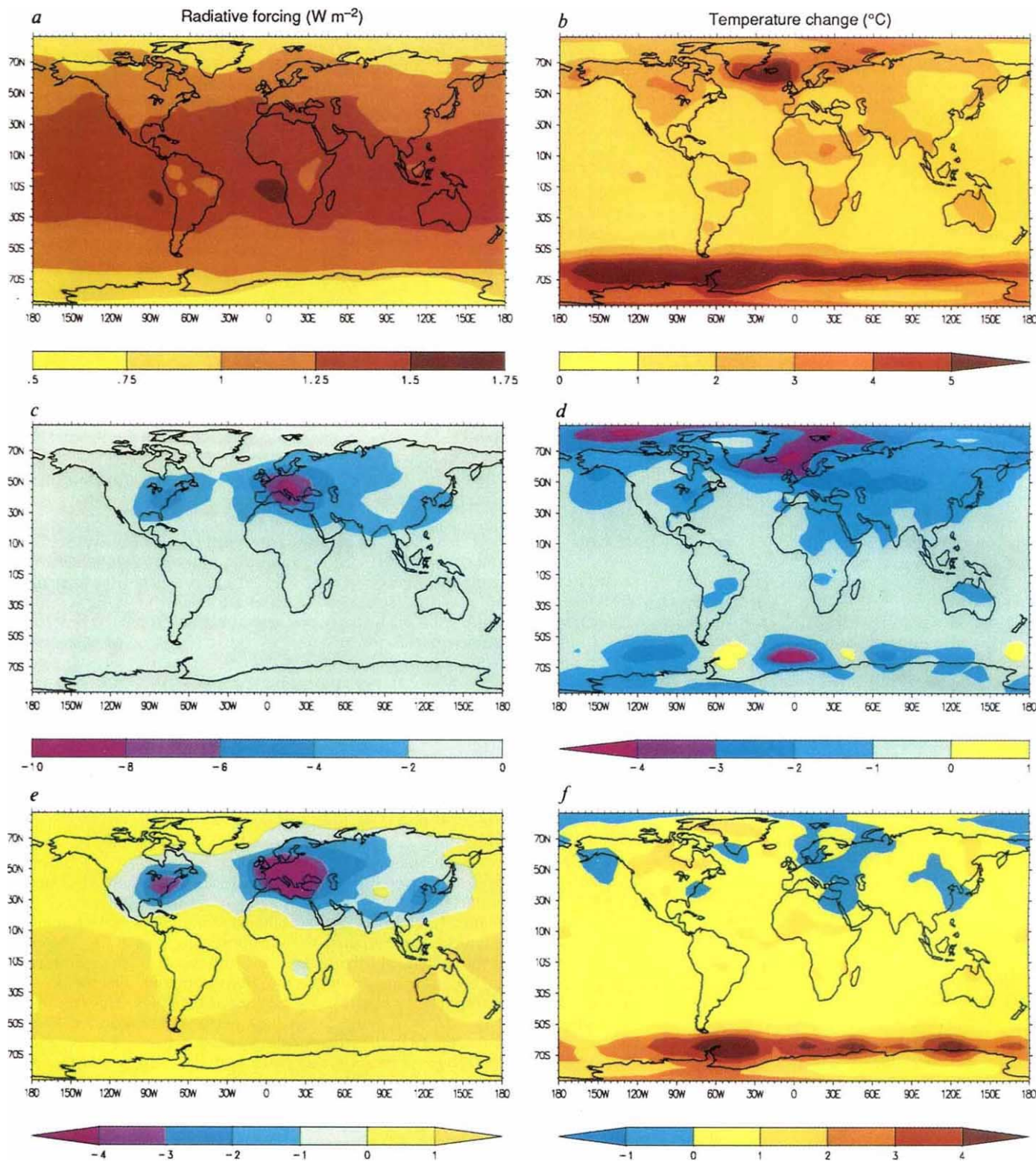


FIG. 1 Climate response to greenhouse forcing, aerosol forcing, and combined greenhouse and aerosol forcing. *a*, The annual average greenhouse forcing (W m^{-2}) due to an increase in atmospheric CO_2 from its pre-industrial value of 275 p.p.m. to 345 p.p.m. *b*, The difference in the annual average top-of-the-atmosphere outgoing shortwave radiation (W m^{-2}) between simulations with and without anthropogenic sulphate aerosols. *c*, The combined radiative forcing (W m^{-2}) due to CO_2 and anthropogenic aerosols (that is, sum of *a* and *b*). *d*, *e*, *f*, The annual average surface temperature change ($^{\circ}\text{C}$) in response to the forcing shown in *a*, *c* and *e*, respectively. The colour bar below each panel indicates the range of numerical values associated with each colour.

METHODS. In simulations that account for anthropogenic aerosols, the 12-h averaged meteorological fields from the GCM are passed to the tropospheric

chemistry model which calculates the sulphate aerosol concentrations. These concentrations are then passed to the GCM for use in the radiative transfer calculations of climate forcing. The LLNL version of the NCAR CCM1 model uses a δ -Eddington model of radiation transport, with one band in the ultraviolet, one in the visible, and five in the near-infrared. Compared to previous versions of the model, the optical depth of high clouds was reduced (through a temperature-dependent cloud-drop number density relationship) to account roughly for the effects of ice crystals. The asymmetry factor for anthropogenic aerosols was set to 0.7 in all wavelength bands, while the scattering coefficient was $8.5 \text{ m}^2 \text{ g}^{-1}$ below $0.9 \mu\text{m}$, and $3.5 \text{ m}^2 \text{ g}^{-1}$ in the infrared wavelengths. The model has 12 levels in the vertical, and a horizontal resolution of $\sim 4.5^{\circ}$ latitude and 7.5° longitude.

(negative) forcing over Europe but is associated with changes in sea ice in the Norwegian and Greenland Seas. A tongue of cold air also extends into the Gulf of Alaska, but this is caused by a change in wind patterns and again is not a direct response to local forcing. Cooling is widespread over the globe, with almost no area of heating. Thus, the pattern of response does not reflect the pattern of the forcing.

Figure 1e and f shows the annual average forcing and temperature change for the simulation with combined CO₂ and anthropogenic sulphate forcing. Note that the extreme change in temperatures at polar latitudes in the Northern Hemisphere has disappeared. Instead, the balance of forcing between CO₂ and anthropogenic sulphate aerosol (see Table 1) has led to a near cancellation of the sea-ice response and enhancement of temperature change at high northern latitudes. Areas of high negative forcing appear to be associated with areas of cooling, though these are often displaced from the area of maximum forcing. The patches of cool air over southern Greenland and the Gulf of Alaska are a result of northerly wind anomalies induced by the aerosol forcing. Overall the atmospheric circulation considerably reduces the regional differences that would result from a local response to forcing. There is, however, a clear indication that in the Northern Hemisphere, the anthropogenic sulphate aerosol generally reduces the magnitude of warming compared to the Southern Hemisphere.

Table 1 gives the temperature response for the three scenarios: CO₂-only, anthropogenic sulphate-aerosol-only, and combined forcing. Note that the combined response is not simply equal to the sum of the individual temperature changes; a linear model would have predicted almost a 40% greater global warming for the combined CO₂-aerosol forcing than our model. It is also important to note, however, that all values in Table 1 are based on 10-year averages (following a 10-year 'spin-up' for each case). The model is known to exhibit interannual variability large enough to make decadal time averages of global mean temperature uncertain to ~0.1 °C. Thus it is not possible to attribute positively the behaviour described above to nonlinearities. Nevertheless, it would not be surprising if nonlinearities associated with sea-ice response, for example, led to a significant nonlinear temperature response.

The climate sensitivity (that is, the ratio of global-average temperature response to global-average forcing) also seems to depend critically on the forcing pattern. For the CO₂-only scenario the sensitivity is 1.7 °C W⁻¹ m², but for the sulphate-only scenario it is only 1.0 °C W⁻¹ m². Furthermore, the sensitivity in both these cases is larger than the estimate quoted earlier (0.76 °C W⁻¹ m²) obtained in fixed sea surface temperature experiments under perpetual July conditions (compare Senior and Mitchell¹⁶). The temperature change of 2.1 °C found in our experiment for an increase of atmospheric CO₂ from 275 p.p.m. to 345 p.p.m. would imply a temperature change of $(2.1 \times \ln 2) / \ln(345/275) = 6.4$ °C for a doubling of CO₂, according to the accepted scaling. Thus, the model exhibits a sensitivity that is at the high end of climate model estimates. Although this high sensitivity is difficult to reconcile with historical data, it cannot be ruled out¹⁸. Further analysis is needed to unravel the reasons for the high model sensitivity, but it may be attributable to such factors as the radiative treatment of high clouds, the upper air cold bias and the sea-ice component of the model. We would like to extend our simulations to verify that our 10-year averages yield reasonable estimates of the sensitivity that would be obtained in longer integrations. This may be important, as a similar model¹⁹ exhibits significant variability on multi-decadal timescales, although the variable-depth mixed layer in that model could introduce low frequency variability that would be absent in the present model. If longer simulations confirm that climate sensitivity depends on the nature of the forcing, then it may be necessary to separate out the individual components of the total forcing to predict the global-mean climate response.

The predicted distinct regional patterns of temperature

response found in our experiments lend credibility to the hope that a signature of the atmosphere's temperature response to anthropogenic forcings may be detectable in the future. Because our predicted pattern of temperature change is clearly a function of the magnitude of the respective positive and negative forcings by greenhouse gases and aerosols and to their spatial distributions, better quantification of these forcings, and in particular the aerosol forcing is needed²⁰. Additionally, similar experiments should be undertaken with other climate models. □

Received 10 January; accepted 18 May 1994.

1. Kiehl, J. T. & Briegleb, B. P. *Science* **260**, 331–314 (1993).
2. Charlson, R. J. et al. *Science* **255**, 423–430 (1992).
3. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
4. Twomey, S. A., Prepgrass, M. & Wolfe, T. L. *Tellus* **36B**, 356–366 (1984).
5. Wigley, T. M. L. *Nature* **339**, 365–367 (1989).
6. Engardt, M. & Rodhe, H. *Geophys. Res. Lett.* **20**, 117–120 (1993).
7. Walton, J. J., MacCracken, M. C. & Ghan, S. J. *J. geophys. Res.* **93**, 8339–8354 (1988).
8. Taylor, K. E. & Ghan, S. J. *J. Clim.* **5**, 907–919 (1992).
9. Covey, C. & Thompson, S. L. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **75**, 331–341 (1989).
10. Cess, R. D. et al. *Science* **245**, 513–516 (1989).
11. Penner, J. E. et al. *J. geophys. Res.* **96**, 959–990 (1991).
12. Penner, J. E., Atherton, C. A. & Graedel, T. in *Proc. 1st Int. Global Atmospheric Biospheric Chemistry* (ed. Prinn, R. G.) (OHIO Conf. Ser., Plenum, New York, 1994).
13. Spiro, P. A., Jacob, D. J. & Logan, J. A. *J. geophys. Res.* **97**, 6023–6036 (1992).
14. Benkovitz, C. *Atmos. Environ.* **16**, 1551–1563 (1982).
15. Charlson, R. J., Langner, J., Rodhe, H., Leovy, C. B. & Warren, S. G. *Tellus* **43AB**, 152–163 (1991).
16. Senior, C. A. & Mitchell, J. F. B. *J. Clim.* **6**, 393–418 (1993).
17. Roeckner, E., Feichter, J. & Siebert, T. *2nd Int. Conf. Modelling Global Climate Change and Variability* (abst.) 78 (Max-Planck-Institut für Meteorologie, Hamburg, 1992).
18. Wigley, T. M. L. & Raper, S. C. B. in *Climate Change: Science, Impacts and Policy* (eds. Jäger, J. & Ferguson, H. L.) 231–242 (Cambridge Univ. Press, 1991).
19. Hansen, J., *J. geophys. Res.* **93**, 9341–9364 (1988).
20. Penner, J. E. et al. *Bull. Am. met. Soc.* **75**, 375–400 (1994).
21. Crutcher, H. L. & Meserve, J. M. NAVAIR 50-1C-52 (Office of Chief of Naval Ops, Washington DC, 1970).
22. Jaeger, L., *Berichte des Deutschen Wetterdienstes* **139**, 1–38 (1976).
23. Rossow, W. B. & Schiffer, R. A. *Bull. Am. met. Soc.* **72**, 2–20 (1991).
24. Gloersen, P. & Campbell, W. J. *J. geophys. Res.* **93**, 10666–10674 (1988).
25. Taljaard, J. J., Van Loon, H., Crutcher, J. L. & Jenne, R. L. NAVAIR 50-1C-55 (Office of Chief of Naval Ops, Washington DC, 1969).

ACKNOWLEDGEMENTS. The comments of T. M. L. Wigley and B. D. Santer are acknowledged. Partial funding for this work was provided by NASA and by the DOE CHAMMP program. Computer time was supplied by a grant from the Model Evaluation Consortium for Climate Assessment. This work was performed under the auspices of the US Department of Energy Environmental Sciences Division by the Lawrence Livermore National Laboratory.

Analogue model of gravitational collapse and surface extension during continental convergence

W. Roger Buck* & Dimitrios Sokoutis†

* Lamont-Doherty Earth Observatory and the Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA

† The Hans Ramberg Tectonic Laboratory, Institute of Earth Sciences, Uppsala University, Norbyvägen 18 B, S-752 36 Uppsala, Sweden

MOST mountain belts occur where continents collide, so it is not surprising that the dominant form of surface deformation is shortening. But there is also abundant geological evidence for extension in the central parts of many mountain belts, at the same time as shortening occurs elsewhere. Previous models for extension require temporal changes in the thermal structure of the lithosphere^{1,2}, the rate of convergence^{2,3}, the strength of the crust³ or the geometry of accretion⁴. Here we present a simple model in which no such changes are required for surface extension during the convergent thickening of a viscous 'crustal' layer. Convergence is driven by the motion of two plates at the base of the layer. In the area where one plate 'subducts' under the other, the surface begins to extend

soon after the start of convergence, eventually stretching by more than 60 per cent. Extension occurs because gravity drives horizontal flow faster at the free surface than in the centre of a viscous layer that is fixed at its base. In real mountain belts, mid-crustal weaknesses may allow the depth-dependent motion required for surface extension.

Within the past ten years it has become accepted that normal faulting is a common, if not ubiquitous, aspect of mountain building. For example, extension of several tens of kilometres occurred during the build-up of the Alps, tectonically exposing rocks of high metamorphic grade⁵. Similar magnitudes of extension occurred in the high Himalaya parallel to the direction of thrusting at lower elevations^{6,7}. Normal faulting is active in the western Apennines⁸ while thrust faulting occurs along the eastern margin⁹. Similarly, normal faulting in the high mountains of the eastern Pamir occurs during overthrusting of the margin of the range¹⁰. Evidence of synchronous normal and thrust faulting is observed in older terrestrial mountain belts such as the Betics of southern Spain⁴. Even on Venus both rift structures and thrusts are imaged in the same highland areas¹¹.

There is great interest in these observations, both because they are unexpected and because they may tell us something about the process of mountain building. To investigate how extension may occur during mountain building we constructed a highly idealized model of crustal deformation. Following others^{3,4}, we assume that an area of continental convergence is analogous to an oceanic subduction zone, except that the low density of crust prevents it from subducting¹². Because crust is weaker than mantle, at typical continental Moho temperatures¹³, the mantle lithosphere is treated as overlapping rigid plates (Fig. 1a).

In our laboratory model (Fig. 1b) a newtonian viscous material takes the place of the crust. We used Rhodorsil Gomme, which has a measured viscosity $\mu = 1.14 \times 10^4$ Pa s, and density $\rho = 1,110$ kg m⁻³. The initial layer thickness h_0 was set at 7 mm and the plate velocity $u_p = 5 \times 10^{-6}$ m s⁻¹, because this gives a surface slope similar to those observed in mountain belts. The

top of the layer was marked with a grid of carbon-black particles using the photocopy method¹⁴ (Fig. 2a).

Figure 2b shows the surface deformation after the plates have converged 25 mm. Extension is indicated where the distance between grid points has increased and the marker circles are stretched into ellipses with long axes parallel to the convergence direction. Surface compression is seen outside this region. Defining strain as the change in grid spacing divided by the initial spacing, the maximum extensional strain seen in this experiment was more than 0.6. The maximum surface slope was of the order of 1/10, so that the error in estimating the surface deformation from the images is less than the error in measuring the grid spacing.

The measured strain at one time in the experiment is plotted in Fig. 3, as is the normalized horizontal surface velocity (u_x/u_p) in the direction of plate motion along the centre of the box. The velocity was calculated using images taken at ~350-s intervals. The surface strain rate is equal to the gradient of u_x with respect to horizontal distance x . A positive gradient indicates extension, and a negative gradient indicates compression. Note that the strain rate is extensional in the region within a layer thickness of the point of convergence ($x=0$ on the plots) and is compressional elsewhere. Figure 4 shows the measured strain at a number of times during this same experiment.

To understand the laboratory results and how they may scale to terrestrial mountain belts, we approximate the flow of the layer using equations for one-dimensional channel flow with a free surface¹⁵. This should be valid as long as the layer is thin compared to the lateral extent of the mountain belt. Temporal changes in the layer thickness are then equal to horizontal gradi-

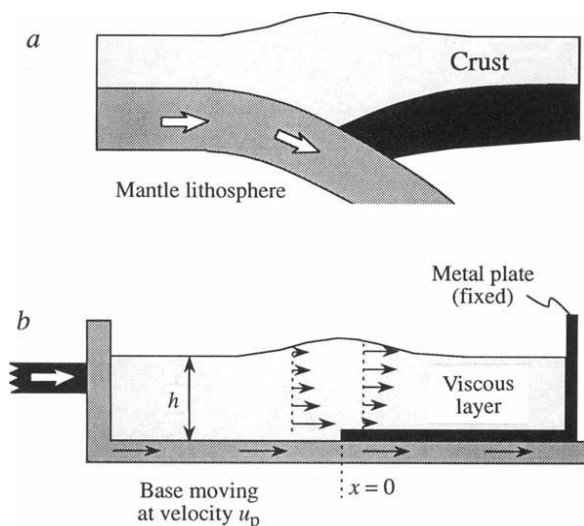


FIG. 1 a, Conceptual model of continental convergence, and b, apparatus for a laboratory model of deformation of a viscous layer. The base of a Plexiglas box is made to move at a constant velocity u_p by an electric motor. The interior of the box is 18 cm long and 7 cm wide. A 0.4-mm-thick metal plate extends 9 cm from the fixed end of the box. The sides of the box were lubricated so that the deformation is nearly two-dimensional. The arrows on b show approximate horizontal velocities which are the sum of plate-driven and gravity-driven flow. The bottom of the layer adheres to the plate and moves with it. Where there is a gradient of layer thickness gravity drives horizontal flow at a rate that increases with distance above the layer base.

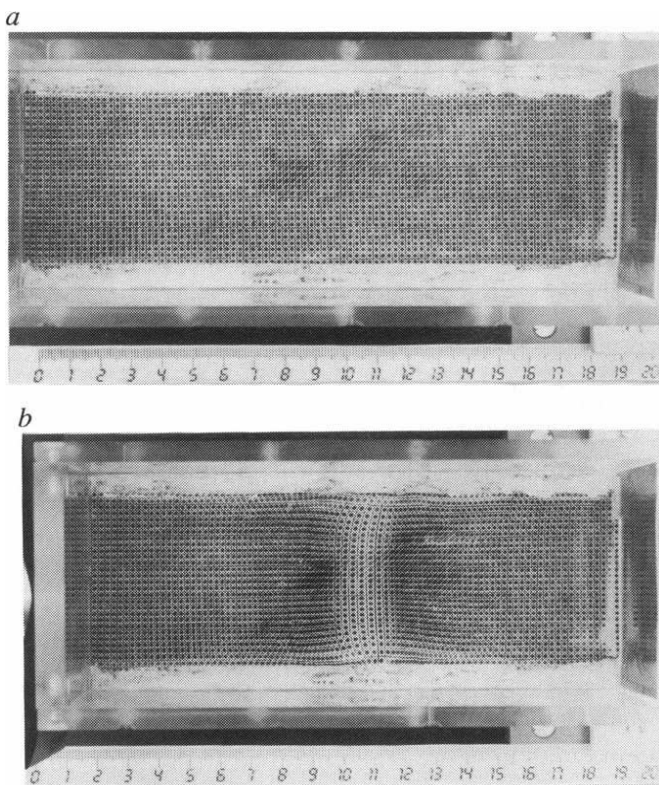


FIG. 2 Photographs of the surface grid on a 7-mm-thick viscous layer. a, Before the start of plate motion, and b after 25 mm convergence at a velocity $u_p = 5 \times 10^{-6}$ m s⁻¹. Extensional strain is seen just downstream (to the right) of the point of basal plate convergence at approximately the 10 cm mark on the scale.

ents in the flux of material driven by the plate and by gravity. This is given by

$$\frac{\partial w}{\partial t} = \frac{\partial}{\partial x} \left(\frac{\rho g h^3}{3\mu} \frac{\partial w}{\partial x} \right) - \frac{\partial}{\partial x} (uh) \quad (1)$$

where $w = h - h_0$, h is the local layer thickness, t is time, g is the acceleration due to gravity and u is the velocity at the base of the layer divided by u_p . Dimensional analysis shows that the Argand number¹⁶ ($Ar = \rho g h^2 / u_p \mu$) controls changes in w .

In the limit that $Ar \gg 1$ we can linearize equation (1) and can neglect the term $\partial(uh)/\partial x$, except where $x=0$, giving

$$\frac{\partial w}{\partial t} = \frac{u_p h Ar}{3} \frac{\partial^2 w}{\partial x^2} \quad (2)$$

At $x=0$ we impose continuity of the flux driven by the plate ($=u_p h_0$) and that driven by gravity. With these limits topography w is symmetric around $x=0$ and we find that $w = (3\pi h_0 u_p t / 4Ar)^{1/2} \text{erfc}(\eta)$, where the similarity variable $\eta = x(3/4Ar u_p t)^{1/2}$ and erfc is the complementary error function. The analytical model mountain grows with time, maintaining an average slope of $\sim 1/(2Ar)$. The surface velocity is the sum of the gravity-driven flow velocity and the velocity of the layer base and is

$$u_s = \begin{cases} u_p [1 - (3/4) \exp(-\eta^2)] & \text{for } x < 0 \\ u_p (3/4) \exp(-\eta^2) & \text{for } x > 0 \end{cases}$$

Thus there is a discontinuity in surface velocity at $x=0$ even though the material flux is continuous.

The amount of surface extension is limited in our model because plate convergence is asymmetric and material moves through a relatively fixed region of extension (Fig. 4). If we assume that the jump in velocity at $x=0$ in our analytical model takes place over a small finite distance δx we can estimate the

surface extension as material moves through that region. The maximum possible extensional strain calculated in that way is 1.1. The key feature of our model leading to surface extension is that the rate of horizontal flow driven by gravity varies with depth.

A finite difference numerical treatment of equation (1) gives a good fit to the analogue model results when we used the measured values of ρ , h_0 , u_p and μ (Figs 3 and 4). To get these fits we had to smooth out the step in basal velocity at $x=0$ over a region just wider than the layer thickness, where two-dimensional flow is significant. The maximum extension is a weak function of the Argand number and so is the mountain slope, showing that the extension is not a result of layer bending.

We tested the effect of isostasy in our numerical model by assuming that the crust floats on the mantle in local isostatic equilibrium (Fig. 1a). To do this, the density ρ is replaced with $\rho^* = (\rho_c / \rho_m)(\rho_m - \rho_c)$ where ρ_c is the crustal density ($2,900 \text{ kg m}^{-3}$) and ρ_m is the mantle density ($3,300 \text{ kg m}^{-3}$). For $h_0 = 35 \text{ km}$, $\mu = 1 \times 10^{21} \text{ Pa s}$ and $u_p = 3 \text{ cm s}^{-1}$ our numerical model gives a slope of $\sim 1/20$, which is the observed topographic slope for the Himalayan front⁹. Several tens of kilometres of surface extension are predicted for this case, which is slightly greater than for a similar case without isostasy.

Using a material with a different rheology in this model should change the amount of surface extension. For a power-law viscous fluid the velocity gradient is proportional to the shear stress raised to the power n . For gravity-driven flow of a layer with $n > 1$ the difference between the surface velocity and the average velocity is less than in the linear ($n=1$) case and so the surface extension is reduced. For $n = \infty$, which is analogous to perfectly plastic flow, our analysis predicts no surface extension. Most previous analytical, numerical and sand-box models of convergence with the same geometry as ours treat the crust as a Coulomb plastic material^{4,17}. These models do not produce surface extension unless temporal changes in material properties are assumed³.

The crust in a mountain belt certainly deforms in a more complex way than a uniform viscous or Coulomb material. However, our analysis of the linear viscous case may provide

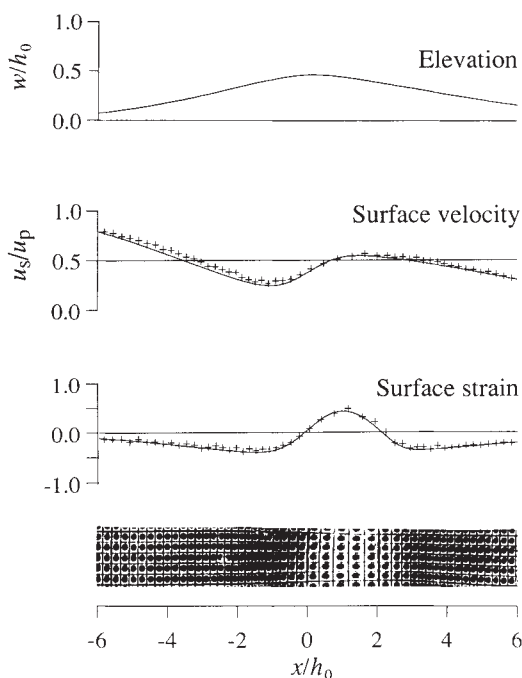


FIG. 3 Detail of the surface grid after convergence of 28 mm ($=4h$) for the same experiment shown in Fig. 2. The horizontal axis is distance from the point of plate convergence x , divided by the initial layer thickness, h_0 . The crosses show the measured surface strain (defined in the text) and surface velocity, u_s , normalized by the plate velocity u_p . The smooth lines show the results of a thin-channel numerical model described in the text, including the predicted surface elevation which was not measured in the laboratory experiment. Elevation is shown in terms of the change in layer thickness, w , normalized by h_0 .

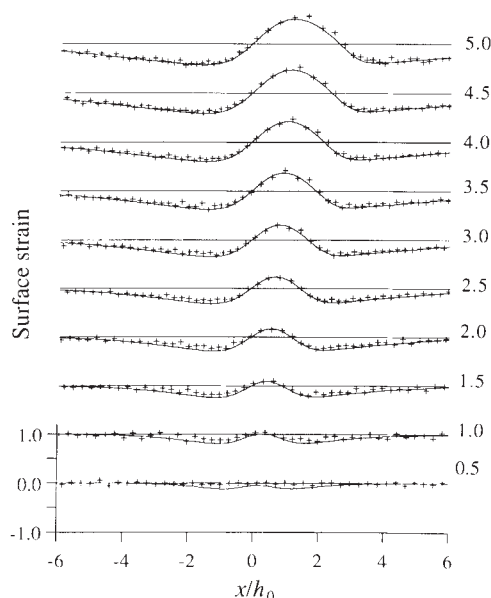


FIG. 4 Surface strain for the laboratory model (crosses) and our numerical model (smooth lines) for a series of time steps. The numbers to the right are time divided by the time it takes for the plate to move a distance h_0 . Extension in the numerical model begins at a non-dimensional model time of ~ 0.7 while the onset in the analogue model is hard to measure. The horizontal axis is labelled as in Fig. 3.

insight into conditions causing surface extension. For example, weak mid-crustal layers or decollement surfaces between stronger crustal layers may develop during orogenesis. A layered rheological structure may allow the kind of depth-variable rates of motion required for extension in our simple viscous model. □

Received 25 January; accepted 9 May 1994.

1. Dewey, J. F. *Tectonics* **7**, 1123–1139 (1988).
2. England, P. C. & Molnar, P. in *Himalayan Tectonics* (eds Treloar, P. J. & Searle, M. P.) 401–411 (Spec. Publ. No. 74, Geol. Soc., London, 1993).
3. Willett, S., Beaumont, C. & Fullsack, P. *Geology* **21**, 371–374 (1993).
4. Platt, J. P. *Geol. Soc. Am. Bull.* **97**, 1037–1053 (1986).
5. Selverstone, J. *Tectonics* **7**, 87–105 (1988).
6. Burg, J. P., Brunel, M., Gapais, D., Chen, G. M. & Liu, G. H. *J. struct. Geol.* **6**, 535–542 (1984).
7. Burchfiel, B. C. & Royden, L. H. *Geology* **13**, 679–682 (1985).
8. Anderson, H. & Jackson, J. *Geophys. J. R. astr. Soc.* **91**, 937–983 (1987).
9. Royden, L. H., Patacca, E. & Scandone, P. *Geology* **15**, 714–717 (1987).
10. Arnaud, N. O., Brunel, M., Cantagrel, J. M. & Tapponier, P. *Tectonics* **12**, 1335–1346 (1993).
11. Solomon, S. C. et al. *Science* **252**, 297–312 (1991).
12. Molnar, P. & Gray, D. *Geology* **7**, 58–62 (1979).
13. Brace, W. F. & Kohlstedt, D. L. *J. geophys. Res.* **85**, 6248–6252 (1981).
14. Dixon, J. M. & Summers, J. M. *J. struct. Geol.* **7**, 83–102 (1985).
15. Turcotte, D. L. & Schubert, G. *Geodynamics* (Wiley, New York, 1982).
16. England, P. C. & McKenzie, D. P. *Geophys. J. R. astr. Soc.* **70**, 295–321 (1982).
17. Byrne, D. E., Wang, W.-h. & Davis, D. M. *Tectonics* **12**, 123–144 (1993).

ACKNOWLEDGEMENTS. We thank P. England, P. Molnar, L. Royden, P. Gautier and C. Talbot for reviews and suggestions. W.R.B. was supported by the US National Science Foundation, and D.S. was supported by the Swedish Natural Science Research Council.

Zinc and carbon co-limitation of marine phytoplankton

F. M. M. Morel, J. R. Reinfelder, S. B. Roberts, C. P. Chamberlain*, J. G. Lee & D. Yee

R. M. Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

* Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

PROCESSES that control carbon uptake by marine phytoplankton are important in the global carbon cycle^{1–3}. Uptake of CO₂ itself may be limited by diffusion⁴. Bicarbonate uptake may be limited by zinc as HCO₃[–] transport appears to involve the zinc metallo-enzyme carbonic anhydrase^{5,6} and the concentration of inorganic zinc in seawater⁷ is low enough to limit the growth of certain phytoplankton in culture^{8,9}. Here we show that HCO₃[–] uptake by the marine diatom *Thalassiosira weissflogii* is modulated by the partial pressure of CO₂ and by the concentration of inorganic Zn (for which Cd and Co may substitute in carbonic anhydrase). This result leads naturally to a 'zinc hypothesis' which, like the standing 'iron hypothesis'¹⁰, posits that Zn (Fe) may limit oceanic production and influence the global carbon cycle. Because of the large ¹³C enrichment of HCO₃[–] over CO₂, our results may be important for the interpretation of δ¹³C measurements in seawater and sediments.

We tested the effects of varying CO₂ and inorganic Zn concentrations on the carbonic anhydrase activity, growth rate and carbon uptake in batch cultures of *T. weissflogii*. The activity of diatom carbonic anhydrase, separated from other proteins by non-denaturing polyacrylamide gel electrophoresis, was high in cultures with low partial pressures of CO₂ (*P*_{CO₂} 100 and 300 p.p.m.) and became undetectable when *P*_{CO₂} was increased to 1,000 p.p.m. (see Fig. 1a). The carbonic anhydrase activity was highest in cultures where *P*_{CO₂} had been lowered from 300 to 100 p.p.m. for 18 h (Fig. 1a, lane 4). Autoradiograms of ⁶⁵Zn in the gel show that Zn—in fact, most of the cellular Zn—

coelutes with carbonic anhydrase and that the ⁶⁵Zn in the bands correlates with carbonic anhydrase activity (Fig. 1b).

At low molar concentrations of inorganic zinc ([Zn]), carbonic anhydrase activity decreased but was restored by the addition of either cadmium or cobalt to the medium (Fig. 1c, d). Additional autoradiograms demonstrate the coelution of Cd and Co with carbonic anhydrase activity. The beneficial effect of Cd and Co on the growth of *T. weissflogii* at limiting Zn concentrations¹¹ thus appears to be due at least partially to the replacement of zinc in carbonic anhydrase.

The regulation of carbonic anhydrase activity by CO₂ concentration implies that the enzyme is important for the acquisition of inorganic carbon. The growth rate of *T. weissflogii* should then become dependent on CO₂ concentration at low Zn, and hence low carbonic anhydrase, concentrations. Indeed, at high

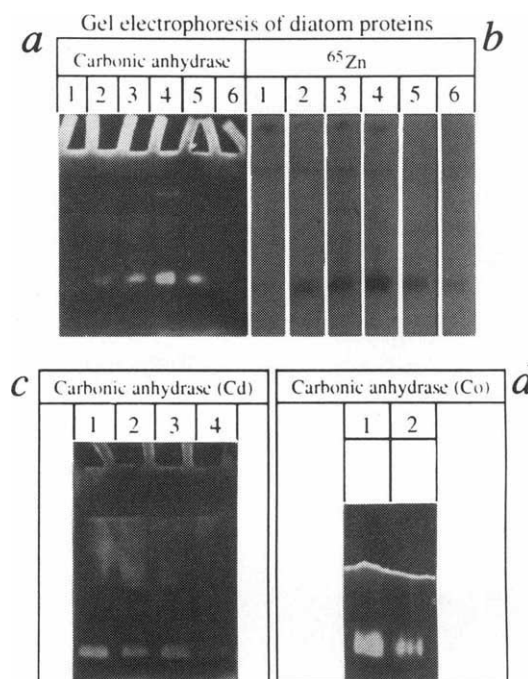


FIG. 1 Carbonic anhydrase and ⁶⁵Zn activity in diatom proteins separated by non-denaturing polyacrylamide gel electrophoresis. a, Carbonic anhydrase activity in *T. weissflogii* grown in Zn-sufficient seawater medium bubbled with air containing 100, 300 or 1,000 p.p.m. CO₂. Lane 1, 1,000 p.p.m.; lane 2, acclimatized at 1,000 p.p.m., switched to 300 p.p.m. for 18 h; lane 3, 300 p.p.m.; lane 4, acclimatized at 300 p.p.m., switched to 100 p.p.m. for 18 h; lane 5, 100 p.p.m.; lane 6, acclimatized at 100 p.p.m., switched to 300 p.p.m. for 18 h. b, Autoradiogram of ⁶⁵Zn activity in gel shown in a. c, Carbonic anhydrase activity in *T. weissflogii* grown in Aquil²⁶ with various molar concentrations of inorganic Cd and Zn. Lane 1, [Cd] = 0.45 nM, [Zn] = 15 pM; lane 2, [Cd] = 0.45 nM, [Zn] = 3.1 pM; lane 3, [Cd] = 9.0 pM, [Zn] = 15 pM; lane 4, [Cd] = 9.0 pM, [Zn] = 3.1 pM. d, Carbonic anhydrase activity in *T. weissflogii* grown with (lane 1) no added Zn, but with [Co] = 21 pM and (lane 2) with no added Co, but with [Zn] = 3.1 pM.

METHODS. *T. weissflogii* cells collected in late-exponential phase were resuspended in 1–2 ml of seawater to yield equal cell densities. The cells were then homogenized by sonication and centrifuged to remove unbroken cells. Equal volumes of lysed cells were loaded in each lane of the 10% polyacrylamide gels which were poured without SDS. SDS and β-mercaptoethanol were also eliminated from the sample buffer. Carbonic anhydrase activity was detected by blowing CO₂ gas on gels soaked in a buffered solution of bromocresol purple (0.1%) as described by Graham et al.²⁷. Bromocresol fluoresces yellow at low pH, and so carbon-anhydrase-catalysed hydration of CO₂ is seen as yellow bands against a purple (pH buffered) background under ultraviolet light. ⁶⁵Zn activity was detected by exposing the gels to X-ray film for 2 weeks.

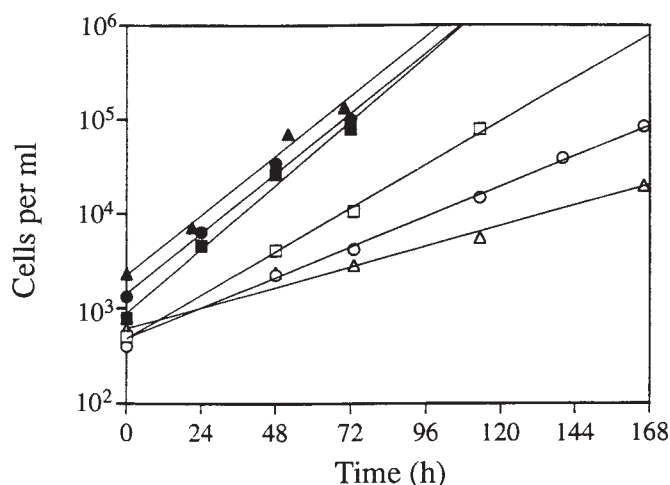


FIG. 2 Growth of *T. weissflogii* in Aquil²⁶ bubbled with air containing 100 p.p.m. (triangles), 300 p.p.m. (circles) or 1,000 p.p.m. (squares) CO_2 and with $[\text{Zn}] = 1.5 \text{ pM}$ (open symbols) or $[\text{Zn}] = 15 \text{ pM}$ (closed symbols).

zinc concentrations ($[\text{Zn}] = 15 \text{ pM}$) the cells grew at their maximum rate, independently of P_{CO_2} , whereas at limiting zinc concentrations ($[\text{Zn}] = 1.5 \text{ pM}$), the growth rate increased with P_{CO_2} (Fig. 2). Presumably the decreased rate of carbonic-anhydrase-facilitated uptake of HCO_3^- at low $[\text{Zn}]$ is partly offset by the availability of CO_2 at high P_{CO_2} .

We used the equilibrium fractionation of ^{13}C between HCO_3^- and CO_2 pools¹² to verify that the effects of $[\text{Zn}]$ and P_{CO_2} on growth rates were due to changes in the cells' ability to take up HCO_3^- and CO_2 . *T. weissflogii* cells grown over a range of CO_2 and inorganic Zn concentrations exhibited varying ratios of stable carbon isotopes. As shown in Fig. 3a, $\delta^{13}\text{C}$ values of cellular carbon decreased dramatically when the P_{CO_2} was

increased from 100 to 1,000 p.p.m. Inverse relationships between the $\delta^{13}\text{C}$ of natural particulate organic carbon and surface P_{CO_2} ranging from 350 p.p.m. down to as low as 200 p.p.m. have been observed^{13,14}. When the activity of carbonic anhydrase was lowered by decreasing $[\text{Zn}]$ in the cultures, the $\delta^{13}\text{C}$ of the cells grown under atmospheric CO_2 also decreased (Fig. 3b). At $[\text{Zn}] = 0.8 \text{ pM}$, the $\delta^{13}\text{C}$ was less than -34‰ , well below the -28‰ we measured when the growth rate was limited to 0.3 divisions per day by lowering the inorganic iron concentration. The very low $\delta^{13}\text{C}$ must thus reflect a specific decrease in HCO_3^- uptake at low $[\text{Zn}]$ rather than slow growth. Stable carbon isotope ratios in seawater and sediments should therefore be interpreted with caution because the $\delta^{13}\text{C}$ of phytoplankton is influenced not only by growth rate, but also by the concentrations of both Zn and CO_2 (which in surface waters ranges from 125 p.p.m.¹⁵ to more than 400 p.p.m.¹⁶).

Because inorganic zinc has been measured at 2 pM in seawater¹⁷, our results lead us to hypothesize that low zinc availability may limit the growth of oceanic phytoplankton. As in the case of the iron hypothesis, it is 'new production' which would presumably be limited by Zn, as large algae, whose low surface area to volume ratio makes them relatively less efficient at taking up Zn (or Fe)^{9,18}, would be most affected⁹. Furthermore, much as Fe limitation seems to be Fe-N co-limitation¹⁹, Zn limitation would be Zn-C co-limitation, and Zn could have a larger impact on the global carbon cycle than Fe. In particular, low zinc concentrations should limit chiefly the growth of organisms that use a carbonic-anhydrase-dependent HCO_3^- transport system, not of coccolithophores, which exhibit little²⁰ or no²¹ carbonic anhydrase activity and can use CO_2 formed from HCO_3^- by the calcification reaction ($2\text{HCO}_3^- + \text{Ca}^{2+} \rightarrow \text{CaCO}_{3(s)} + \text{CO}_2 + \text{H}_2\text{O}$)²². Thus, low zinc concentrations would promote both low productivity and high rain ratios ($\text{CaCO}_3/\text{organic C}$ export) and make the biological carbon pump particularly inefficient, as CaCO_3 export regenerates CO_2 . Conversely, the high Zn (and Fe) inputs to the open oceans resulting from the high levels of terrigenous dust during glacial times^{23,24} may have resulted in a low rain ratio as well as increased productivity. Reconstructed rain ratios from ice age sediments are indeed quite

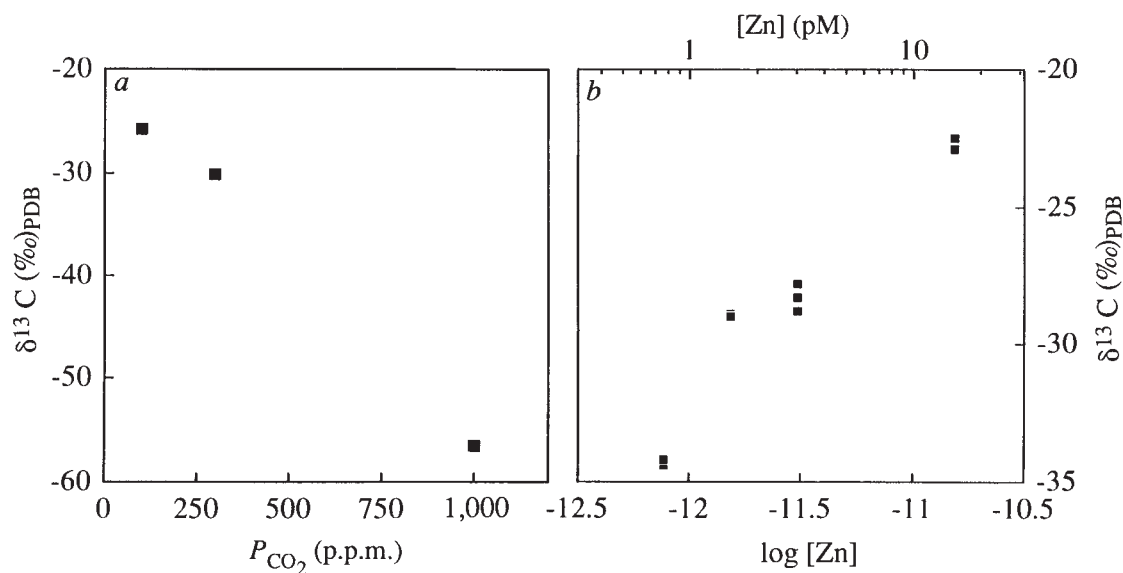


FIG. 3 a, $\delta^{13}\text{C}$ of *T. weissflogii* as a function of P_{CO_2} in the culture medium. Data were calculated according to $\delta^{13}\text{C} = [({}^{13}\text{C}/{}^{12}\text{C}_{\text{sample}}/{}^{13}\text{C}/{}^{12}\text{C}_{\text{ref}} - 1) \times 1,000]$ with respect to PDB standard calcium carbonate. b, $\delta^{13}\text{C}$ of *T. weissflogii* as a function of inorganic Zn concentration in the culture medium. Growth rates for all points in a were ~ 2 divisions per

day. Growth rates in b were 1.7, 0.93, 0.59 and 0.32 divisions per day for $[\text{Zn}] = 15, 3.1, 1.5$ and 0.77 . The $\delta^{13}\text{C}$ values of phytoplankton grown at different concentrations of CO_2 (a) are artificially low through the use of a CO_2 source that was $\sim 20\text{‰}$ lighter than atmospheric CO_2 , which was used for the experiments at varying $[\text{Zn}]$ (b).

low²⁵. The zinc and the Fe hypotheses may thus explain the low atmospheric CO₂ during the last glacial maximum. □

Received 14 February; accepted 1 June 1994.

1. Raven, J. A. & Johnston, A. M. *Limnol. Oceanogr.* **36**, 1701–1714 (1991).
2. Broecker, W. S. *Globi Biogeochem. Cycles* **5**, 191–192 (1991).
3. Sarmiento, J. L. & Toggweiler, J. R. *Nature* **308**, 621–624 (1984).
4. Riebesell, U., Wolf-Gladrow, D. A. & Smetacek, V. S. *Nature* **361**, 249–251 (1993).
5. Sultemeyer, D. F., Miller, A. G., Espie, G. S., Fock, H. P. & Canvin, D. T. *Pl. Physiol.* **89**, 1213–1219 (1989).
6. Price, G. D. & Badger, M. R. *Pl. Physiol.* **89**, 37–43 (1989).
7. Bruland, K. W. *Limnol. Oceanogr.* **34**, 269–285 (1989).
8. Anderson, M. A., Morel, F. M. M. & Guillard, R. R. L. *Nature* **276**, 70–71 (1978).
9. Sunda, W. G. & Huntsman, S. A. *Limnol. Oceanogr.* **37**, 25–40 (1992).
10. Martin, J. H. & Fitzwater, S. E. *Nature* **331**, 341–343 (1988).
11. Price, N. M. & Morel, F. M. M. *Nature* **344**, 658–660 (1990).

12. Mook, W. G., Bommerson, J. C. & Staverman, W. H. *Earth planet. Sci. Lett.* **22**, 169–176 (1974).
13. Rau, G. H., Takahashi, T., Des Marais, D. J., Repeta, D. J. & Martin, J. H. *Geochim. cosmochim. Acta* **56**, 1413–1419 (1992).
14. Freeman, K. H. & Hayes, J. M. *Globi Biogeochem. Cycles* **6**, 185–198 (1992).
15. Codispoti, L. A., Friedrich, G. E., Iverson, R. L. & Hood, D. W. *Nature* **296**, 242–245 (1982).
16. Broecker, W. S. & Peng, T.-H. *Tracers in the Sea* (Eldigio, Palisades, New York, 1982).
17. Bruland, K. W. *Limnol. Oceanogr.* **37**, 1008–1017 (1992).
18. Hudson, R. J. M. & Morel, F. M. M. *Deep-Sea Res.* **40**, 129–150 (1993).
19. Price, N. M., Ahner, B. A. & Morel, F. M. M. *Limnol. Oceanogr.* **39**, 520–534 (1994).
20. Quiroga, O. & Gonzalez, E. L. *J. Phycol.* **29**, 321–324 (1993).
21. Sikes, S. S. & Wheeler, A. P. *J. Phycol.* **18**, 423–426 (1982).
22. Nimer, N. A. & Merrett, M. J. *New Phytol.* **121**, 173–177 (1992).
23. Petit, J. R., Briat, M. & Royer, A. *Nature* **293**, 391–393 (1981).
24. Arimoto, R., Duce, R. A., Ray, B. J., Hewitt, A. B. & Williams, J. J. *geophys. Res.* **92**, 8465–8479 (1987).
25. Pedersen, T. F. *Geology* **11**, 16–19 (1983).
26. Price, N. M. et al. *Biol. Oceanogr.* **6**, 443–461 (1988/89).
27. Graham, D., Reed, M. L., Patterson, B. D. & Hockley, D. G. *Annls N.Y. Acad. Sci.* **429**, 222–237 (1984).

Where we look when we steer

M. F. Land* & D. N. Lee†

*Sussex Centre for Neuroscience, School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK

†Perception in Action Laboratories, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK

STEERING a car requires visual information from the changing pattern of the road ahead. There are many theories about what features a driver might use^{1–3}, and recent attempts to engineer self-steering vehicles have sharpened interest in the mechanisms involved^{4,5}. However, there is little direct information linking steer-

ing performance to the driver's direction of gaze³. We have made simultaneous recordings of steering-wheel angle and drivers' gaze direction during a series of drives along a tortuous road. We found that drivers rely particularly on the 'tangent point' on the inside of each curve, seeking this point 1–2 s before each bend and returning to it throughout the bend. The direction of this point relative to the car's heading predicts the curvature of the road ahead, and we examine the way this information is used.

Steering performance was measured using a Jaguar car equipped with instruments to record and store information on the angle of the steering wheel, speed and other parameters. The car had automatic transmission and power steering. Gaze direction was measured using a head-based video system which imaged both the road ahead and the driver's eye^{6,7}. A computer

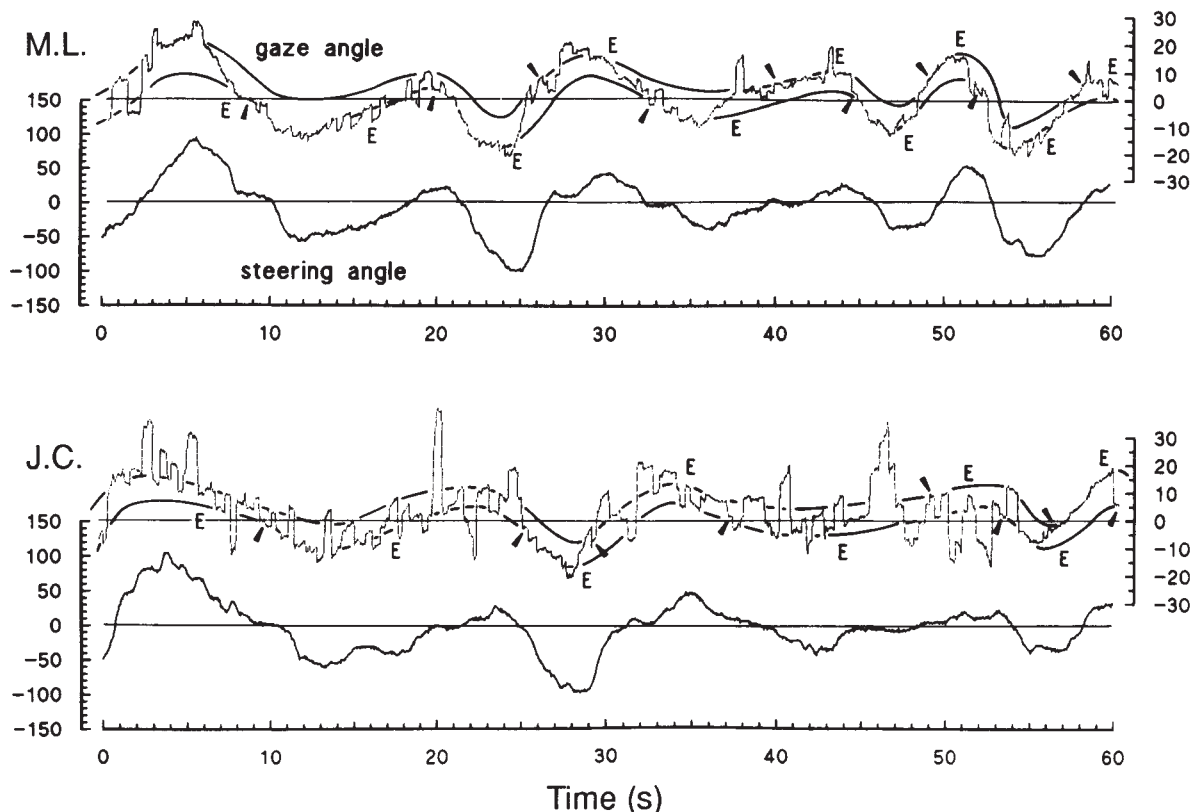
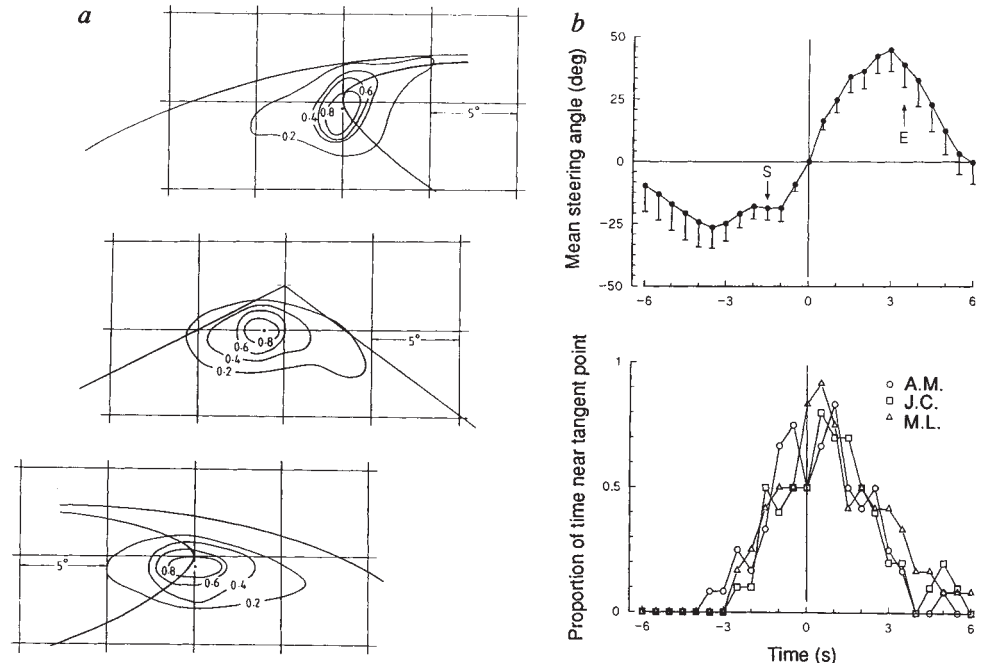


FIG. 1 Simultaneous records of gaze angle relative to the car's heading and steering-wheel angle for two drivers (M.L. and J.C.) driving the same twisting road. M.L. drove slightly faster (45 km h⁻¹) than J.C. (40 km h⁻¹). The performance of a third driver (A.M.) was very similar to M.L.'s. The approximate locations of the edges of the road have been added to the gaze angle traces. This was done by measuring on the video the positions of the sides of the road at the distance at which gaze direction intersected the road (typically 20–30 m). These records show how gaze moves to the right road edge on right-hand bends and

to the left road edge on left-hand bends. Arrowheads indicate the first clear saccade to the tangent point preceding each bend, and E is the point on the exit from each bend where the tangent point disappears. J.C. spends more time looking off the road than M.L., but still returns frequently to the tangent points. These may be tracked by repeated refixations (for example, J.C. 9–17 s) or smooth tracking (M.L., 48–51 s, J.C., 56–59 s), or a combination. Records were synchronized to better than 0.1 s by comparing the steering-angle record with steering-wheel movements seen on the gaze video. Both scales are in degrees.

FIG. 2 a, Contour plots showing the distribution of the centre points of all fixations made during right bends (top), left bends (bottom) and when no tangent point was visible (centre). Bends are taken to be between the first fixation on the tangent point (arrowheads in Fig. 1, and S in b) and the time of its disappearance (E in b). Through the bends, all measurements were made with respect to the tangent point itself, and on the straight parts with respect to the vanishing point. Contour intervals are proportions of the maximum value, the absolute value of which is ~ 0.12 fixations per $\text{deg}^2 \text{s}^{-1}$. About 35% of all fixations lie beyond the 0.2 contour, and are widely spread. Figures have been positioned so that the car's heading is aligned on all three, and corresponds to the centre line in the centre figure. Data from the drivers have been pooled; all three drivers showed very similar distributions on the road itself, although J.C. made relatively more remote fixations. b, Time course of the use of the tangent point. Top, steering-wheel angle adopted around a 'typical' bend averaged from the 10 largest bends on the drive, synchronized to the point when the steering wheel angle is zero (0 s on x-axis). Bars indicate 1 s.e. Arrow at S shows average timing of the first saccade to tangent point, and E the average timing of its disappearance. Bottom,



model of the eye was used to convert the recorded position of the iris into direction of view relative to the head, and this was added as a spot to the image of the road. Head movements were monitored from the movement on the video image of a series of markers on the car's windscreen. Eye-in-head and head-in-car coordinates were added to give eye-in-car (gaze) records (Fig. 1). Gaze measurements are repeatable to better than 1° , and the absolute accuracy in the central 20° region is about 1° . Three experienced drivers (J.C., A.M. and M.L.) drove at normal speeds around Arthur's Seat in Edinburgh. This road has many bends, but is single-lane and one way and so provides demanding steering conditions uncomplicated by traffic.

Figure 1 shows 1-min records of steering-wheel angle and gaze direction from two drivers. Both gaze records show a typical pattern of saccades, fixations, and smooth tracking⁸. For driver M.L. steering and gaze records are very similar, with the steering wheel turned at an angle corresponding to the direction of gaze relative to the car's heading, after a delay of about 0.75 s. For driver J.C., this relation is less obvious because he spends more time looking off the road at the scenery but, like driver M.L., he keeps returning his gaze to the sides of the road.

Viewing the videos of all three drivers, one is immediately struck by the way the eye seeks and returns to the 'tangent point' (reversal or extremal point) on the inside of each bend, where the edge of the road reverses direction. The drivers themselves were surprised by their consistent use of this point. Figure 2a shows the relative numbers of fixations in the regions around the tangent points. The distributions are centred within a degree of the tangent point itself on both left and right bends. As with straight driving (centre), they are elongated horizontally, with roughly exponential decline from the central peak. The time distribution of these fixations is shown in Fig. 2b. Gaze is directed to the tangent point, with a rather obvious saccade 1–2 s before the car enters each bend, and remains there with relatively few excursions for ~ 3 s into the bend. Half a second after the car has entered the bend, the gaze of all three drivers is directed to the tangent point for about 80% of the time (Fig.

the portion of the total time spent by the gaze of each driver within 3° of the tangent point. It can be seen that during the first second into the turn, all three drivers look at the tangent point almost all ($>75\%$) of the time.

2b). Clearly, the tangent point is visually important to the driver when steering into a turn.

Why should this be so? A likely answer is that tangent point direction (θ) relative to the car's heading is a very good predictor of the curvature of the road ahead². It is easy to show that the curvature of the road between the car and tangent point is related to θ by

$$C = 1/(d \cos \theta) - 1/d \quad (1)$$

where the curvature C is the reciprocal of the local inner radius of the road, and d is the lateral distance of the driver from the kerb. If d can be maintained by using some independent measurement, for example the perspective angle made by the kerb or lane marking near the car, then C is easily obtained from θ . And as the steering-wheel angle is directly proportional to C , equation (1) provides a simple steering strategy. In other schemes, the relative motion of the tangent point ($d\theta/dt$) can be used either to remove the need to compute d separately² or to monitor departures from a predicted constant curvature path⁹. Such predictions might come from the overall velocity flow field, which can specify the car's heading around a curve¹⁰.

On a narrow undulating road, the tangent point is almost the only reliable cue to curvature because, unlike the general shape of the road ahead, it is unaffected by slope changes¹¹. It is possible to steer while deliberately looking at the opposite side of the road, but it feels wrong, and one has to drive more slowly. Here as in other tasks¹², the brain prefers a 'do it where you look' strategy, in which the object guiding a control strategy is placed close to the centre of vision. In more relaxed wide-road driving, much more time is spent on tasks irrelevant to steering, which is why the importance of the tangent point has not previously become apparent³. The performance of driver J.C. (Fig. 1) shows that steering control can 'time-share' with other activities. In urban driving, where one may be called upon simultaneously to steer a course, avoid other vehicles and look for a street name, this capacity for dividing time while avoiding crosstalk between tasks is crucial. □

Received 24 March; accepted 20 May 1994.

1. Riemersma, J. B. J. in *Vision in Vehicles III* (ed. Gale, A. G.) 163–170 (North-Holland, Amsterdam, 1991).
2. Raviv, D. & Herman, M. in *Proc. IEEE Workshop on Visual Motion, Princeton, NJ* 217–225 (1991).
3. Serafin, C. Univ. Michigan Transportation Res. Lab. Document UMTRI-93-29 1–63 (1993).
4. Okuno, A., Fujita, K. & Kutami, A. in *Vision-based Vehicle Guidance* (ed. Masaki, I.) 222–237 (Springer, New York, 1992).
5. Dickmanns, E. D. et al. in *Proc. IEEE 4th Intl. Conf. on Computer Vision, Berlin* 608–615 (1993).
6. Land, M. F. *Nature* **359**, 318–320 (1992).
7. Land, M. F. in *Proc. IEEE Systems, Man and Cybernetics Conf., Le Touquet 1993* Vol. 3, 490–494 (1993).
8. Carpenter, R. H. S. *Movements of the Eyes* (Pion, London, 1988).
9. Lee, D. N. & Lishman, J. R. *Scand. J. Psychol.* **18**, 224–230 (1977).
10. Warren, W. H. Jr, Mestre, D. R., Blackwell, A. W. & Marris, M. W. J. *exp. Psychol.* **17**, 28–43 (1991).
11. Landwehr, K. in *Vision in Vehicles III* (ed. Gale, A. G.) 187–194 North-Holland, Amsterdam, 1991).
12. Ballard, D. H., Hayhoe, M. M., Li, F. & Whitehead, S. D. *Phil. Trans. R. Soc. B* **337**, 331–339 (1992).

ACKNOWLEDGEMENTS. We thank R. Wakeling of Ford UK for arranging the loan of the car, Y. Coello and J. Horwood for technical help, J. Cutburt and A. Marshall for driving the car (the third driver was M.F.L.), and T. Collett for reading and commenting on the manuscript. This study was supported in part by a grant from the Joint Council Initiative on Cognitive Science and Human Computer Interaction, and by an SERC grant to the Sussex Centre for Neuroscience.

Glutamate-mediated astrocyte–neuron signalling

Vladimir Parpura*†, Trent A. Basarsky*†, Fang Liu†‡, Ksenija Jeftinija†‡, Srdija Jeftinija†‡ & Philip G. Haydon*†§

* Department of Zoology and Genetics, † Neuroscience Program, or ‡ Department of Veterinary Anatomy, Iowa State University, 339 Science II, Ames, Iowa 50011, USA

§ To whom correspondence should be addressed

NEUROTRANSMITTER released from neurons is known to signal to neighbouring neurons and glia^{1–3}. Here we demonstrate an additional signalling pathway in which glutamate is released from astrocytes and causes an NMDA (*N*-methyl-D-aspartate) receptor-mediated increase in neuronal calcium. Internal calcium was elevated

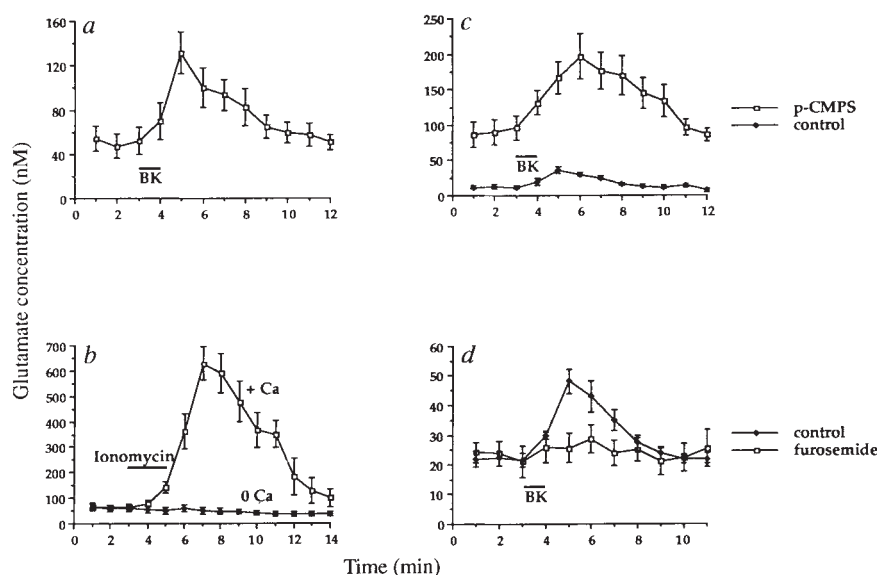
and glutamate release stimulated by application of the neuro-ligand bradykinin to cultured astrocytes. Elevation of astrocyte internal calcium was also sufficient to induce glutamate release. To determine whether this released glutamate signals to neurons, we studied astrocyte–neuron co-cultures. Bradykinin significantly increased calcium levels in neurons co-cultured with astrocytes, but not in solitary neurons. The glutamate receptor antagonists D-2-amino-5-phosphonopentanoic acid and D-glutamylglycine prevented bradykinin-induced neuronal calcium elevation. When single astrocytes were directly stimulated to increase internal calcium and release glutamate, calcium levels of adjacent neurons were increased; this increase could be blocked by D-glutamylglycine. Thus, astrocytes regulate neuronal calcium levels through the calcium-dependent release of glutamate.

The release of glutamate from neuron-free cultures of neocortical astrocytes was monitored using high-performance liquid chromatography (HPLC). The neuro-ligand bradykinin (100 nM) increased glutamate in the superfusate ($n=16$; $P<0.02$, paired *t*-test; Fig. 1a). To investigate whether internal calcium concentration ($[Ca^{2+}]_i$) influences glutamate release from astrocytes, we monitored internal calcium using Fura-2 and found that bradykinin elevated astrocyte calcium concentration from 99 ± 10 to 568 ± 89 nM (mean $\pm$ s.e.m., $n=20$; Fig. 2). We used the calcium ionophore ionomycin to determine whether elevated internal calcium is sufficient to stimulate glutamate release. Addition of ionomycin (5 μ M) in the presence of external calcium, but not in its absence, stimulated release of glutamate from astrocytes (Fig. 1b). Furthermore, bradykinin did not cause significant glutamate release when calcium was removed from the external saline ($P>0.1$, paired *t*-test). These results show that bradykinin induces the release of glutamate from neocortical astrocytes and that elevated internal calcium can induce glutamate release.

The role of glutamate transporters as mediators of bradykinin-induced glutamate release was investigated using glutamate transport inhibitors. Consistent with previous observations^{4,5}, *p*-chloromercuriphenylsulphonic acid (*p*-CMPS; 50 μ M) and *L*-trans-pyrrolidine-2,4-dicarboxylate (PDC; 100 μ M–1 mM) raised the basal level of glutamate in the astrocyte superfusate ($P<0.02$, Mann–Whitney *U*-test⁶). Furthermore, neither

FIG. 1 Bradykinin causes calcium-dependent release of glutamate from astrocytes. The superfusate from astrocyte cultures was collected at 1-min intervals and levels of glutamate were measured using HPLC. a, Bradykinin (100 nM) increases glutamate release from astrocytes ($n=4$). b, Ionomycin (5 μ M) stimulates glutamate release in cultures in calcium-containing saline (+Ca; $n=6$), but not in its absence (0 Ca; $n=7$). c, The glutamate-transport inhibitor *p*-CMPS (50 μ M) raised the basal level of glutamate and enhanced the bradykinin-induced elevation of glutamate in superfusate ($n=6$), presumably as a result of inhibition of glutamate uptake. d, Furosemide (5 mM) blocks bradykinin-induced release of glutamate ($n=4$) without significantly affecting bradykinin-induced astrocyte calcium mobilization.

METHODS. Cultures from 1–4-day-old Sprague–Dawley rat cortices were enriched in type-1 astrocytes according to ref. 25 and were maintained in α -MEM medium. These cultures were neuron-free, as revealed by immunocytochemistry using an antibody directed against the synaptic protein synaptotagmin (1:250; clone 41.1, provided by R. Jahn). The amino-acid content of samples was determined by HPLC with fluorescence detection. Before injection, aliquots of samples were derivatized with *o*-phthalic aldehyde (OPA) 2-mercaptoethanol reagent (Pierce). Chromatography was performed on a 15 cm Microsorb-MV HPLC column (Rainin Instrument Co.) using a sodium acetate (pH 5.9)–



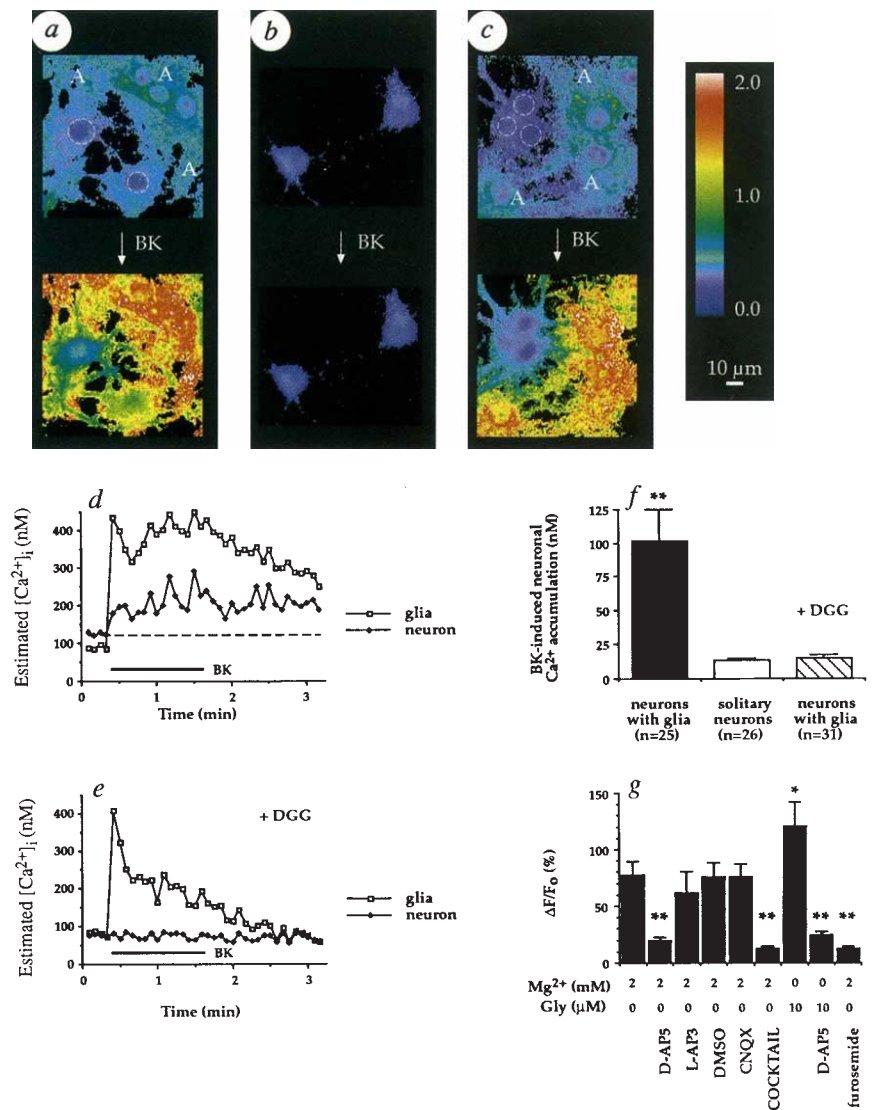
methanol gradient. Points represent mean $\pm$ s.e.m. Solutions: A modified Ringer's solution used for perfusion contained (in mM): NaCl 128, KCl 1.9, KH_2PO_4 1.2, $CaCl_2$ 2.4, $MgSO_4$ 1.3, $NaHCO_3$ 26, and glucose 10 (pH=7.4). In 'zero Ca' solution, calcium was replaced by 1 mM EGTA, 2.5 mM $MgSO_4$, 0.2 mM $CaCl_2$ to yield 24 nM free calcium.

FIG. 2 Bradykinin causes a glutamate-mediated elevation of neuronal calcium level in neurons co-cultured with astrocytes. Calcium levels in neocortical neurons and astrocytes were monitored using Fura-2. **a**, Puffer application of bradykinin (BK; 1 μ M, 75 s) elevated calcium levels of neurons (circled) and astrocytes (A). When neurons were cultured alone (**b**), bradykinin did not significantly alter their calcium levels. In **c**, three neurons are shown on top of astrocytes in the presence of DGG (1 mM). Bradykinin elevated the astrocyte calcium levels, but not the level of calcium in adjacent neurons. Colour scale is a linear pseudocolour representation of $[Ca^{2+}]_i$, shown in the form of the ratio of Fura-2 emission at 510 nm due to sequential excitation at 350 and 380 nm. Scale bar, 10 μ m. **d** and **e**, Time course of astrocyte and neuronal calcium responses after addition of bradykinin in control conditions (**d**) and in the presence of DGG (1 mM) (**e**). The astrocyte change in calcium (peak value subtracted from resting) induced by bradykinin (469 ± 86 nM, $n=20$; control) was not significantly affected by DGG (424 ± 45 nM, $n=22$; Student's *t*-test, $P>0.6$). In **f**, the mean bradykinin-evoked calcium accumulations from experiments reported in **a–e** are shown. **g**, Calcium elevations in neurons co-cultured with astrocytes were monitored using Fluo-3 and confocal microscopy. Changes in calcium are represented as $\Delta F/F_0$. Bradykinin-induced elevations of neuronal calcium were reduced by D-AP5 (50 μ M). Mg^{2+} -deficient medium containing glycine (10 μ M) augmented the neuronal calcium response to bradykinin, which was significantly reduced by D-AP5 (50 μ M). Furosemide (5 mM) reduced neuronal calcium elevation in response to bradykinin. Bars represent means $\pm$ s.e.m. ($n=18$ for each group in **g**).

METHODS. Dissociated cultures from visual cortices were prepared using a modification of the Heutner & Baughman procedure^{26–28} and were used after 10–15 days in culture. A monoclonal antibody directed against glial fibrillary acidic protein (GFAP, 1:500; ICN Immunobiologicals), showed that glial cells in this culture were astrocytes ($n=159$ of 159 tested). Calcium imaging: **a–f**, Cells were loaded with Fura-2 AM (2 μ M) and calcium levels were estimated according to ref. 28. In **g**, cells were loaded with Fluo-3AM (20 μ g ml⁻¹)²⁹. Images were acquired using a Noran Odyssey real-time confocal microscope (Noran). Imaging saline contained (in mM): NaCl 140, KCl 5, $MgCl_2$ 2, $CaCl_2$ 2 and HEPES 10 (pH 7.35). In magnesium-deficient saline, Mg^{2+} was replaced by calcium and glycine

p-CMPS (Fig. 1c, $n=6$) nor PDC ($n=6$) impaired the ability of bradykinin to stimulate glutamate release from astrocytes. Although glutamate transporter inhibitors block glutamate uptake, it is not known whether they also block reversed uptake. Furosemide, an anion-transport inhibitor, reduces the hypo-osmotic-induced release of glutamate from astrocytes⁷. Addition of furosemide (5 mM) to cortical astrocytes blocks bradykinin-induced release of glutamate (Fig. 1d) compared with controls ($P<0.05$, Mann–Whitney *U*-test) without significantly affecting bradykinin's calcium-mobilizing action (bradykinin-induced calcium mobilization in furosemide was 714 ± 56 nM, $n=29$, compared with 609 ± 107 nM in control, $n=28$; $P>0.3$ Student's *t*-test). Thus, bradykinin mobilizes calcium to stimulate glutamate release through a furosemide-sensitive release mechanism.

To see whether glutamate released from astrocytes acts on adjacent neurons, we studied the effect of bradykinin on $[Ca^{2+}]_i$ in neurons co-cultured with astrocytes. Bradykinin causes an increase in neuronal calcium only when neurons were in contact with astrocytes (Fig. 2). Bradykinin induced a neuronal calcium accumulation of 101 ± 25 nM ($n=25$; Fig. 2a, d, f). When bradykinin was applied, calcium levels increased significantly



(10 μ M) was added. Pharmacological agents were bath applied 10 min before application of bradykinin. Cocktail contained D-AP5 (50 μ M), CNQX (10 μ M) and L-AP3 (1 mM). Statistical analysis: One-way ANOVA followed by Scheffé's (*f*) or Tukey's (*g*) *post hoc* comparisons was used. Significance was established at $P<0.05$ (*) and $P<0.01$ (**).

in astrocytes ($n=20$) and neurons ($n=25$) after 15 and 50 s, respectively (Tukey's *post hoc*, $P<0.05$). When neurons were cultured alone, they did not respond significantly to bradykinin (calcium accumulation of 13 ± 2 nM, $n=26$; Fig. 2b, f). To determine whether bradykinin acts directly on neurons to elevate neuronal calcium levels or acts indirectly through glutamate released from astrocytes, we added a broad-spectrum glutamate receptor antagonist, D-glutamylglycine⁸ (DGG). DGG (1 mM) significantly reduced bradykinin-induced calcium accumulation in neurons to 15 ± 2 nM ($n=31$), as compared to 101 ± 25 nM in parallel cultures (Fig. 2c, e, f), without altering the astrocyte response to bradykinin (Fig. 2e). DGG was washed out of the bath and glutamate was applied to neurons to determine whether glutamate receptors were present. Direct application of glutamate (100 μ M) elevated internal calcium in each category of neurons shown in Fig. 2a–c and f (neurons and glia, solitary neurons, and neurons and glia previously exposed to DGG). These responses were not significantly different from one another (Scheffé's *post hoc* comparison⁹). These data suggest that bradykinin elevates neuronal calcium by the action of glutamate, which is released from astrocytes in response to bradykinin.

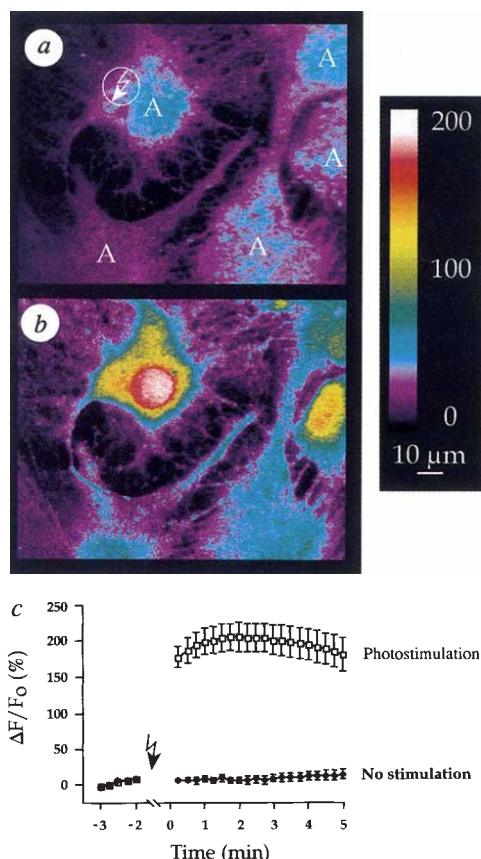


FIG. 3 Photostimulation of a single astrocyte causes an increase in $[Ca^{2+}]_i$ in an astrocyte (A) network. An astrocyte in a was exposed to UV light (2 min) from a xenon-arc lamp which caused a sustained increase in $[Ca^{2+}]_i$ in this cell and in unstimulated neighbours (b). A circled lightning bolt represents the location and diameter of UV exposure. Photostimulation elevated calcium for several minutes in all astrocytes in the field of interest without affecting their morphology or causing a leakage of cytoplasmic components, as judged by the maintenance of the calcium indicator Fluo-3. Summary of data is shown in c. Colour scale indicates linear pseudocolour representation of fluorescence intensity ranging from 0 to 200 units. Points represent mean $\pm$ s.e.m. ($n = 7$ for stimulated, and $n = 15$ for control group). Scale bar, 10 μ m.

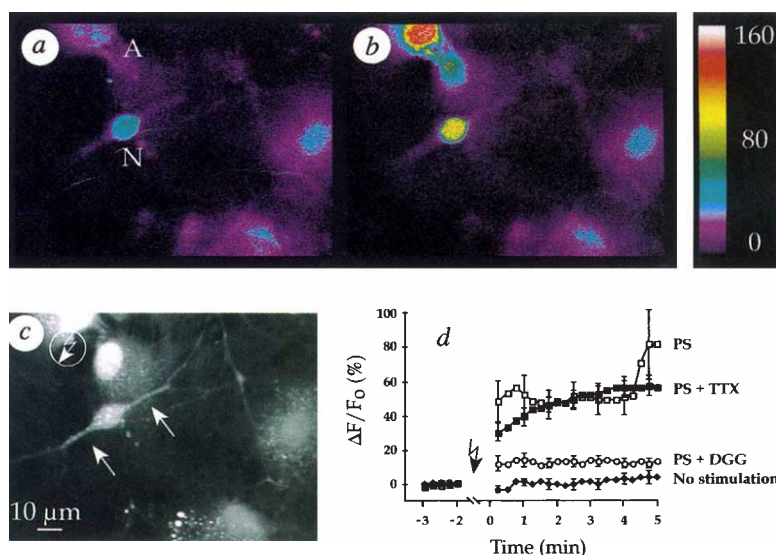
not 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX; 10 μ M) or L-2-amino-3-phosphonopropionic acid (L-AP3; 1 mM), significantly reduced the increase in bradykinin-induced neuronal calcium (Fig. 2g). Consistent with a role for NMDA receptors in mediating the action of released glutamate, removal of external Mg^{2+} and addition of glycine (10 μ M) significantly augmented the bradykinin-induced neuronal calcium elevation (Fig. 2g).

HPLC showed that furosemide blocks bradykinin-induced glutamate release from astrocytes (Fig. 1). If glutamate released from astrocytes signals to adjacent neurons, then furosemide should block this intercellular signalling pathway. Furosemide (5 mM) significantly reduced the ability of bradykinin to raise neuronal calcium (by 84%; Fig. 2g) and reduced the increase in neuronal calcium in response to focally applied glutamate (100 μ M) (by 43%; $n = 25$; $P < 0.001$, Student's *t*-test). As furosemide blocks bradykinin-induced glutamate release from astrocytes and reduces glutamate-induced neuronal calcium mobilization, the sensitivity to furosemide of the bradykinin-induced neuronal calcium elevation is consistent with the mediation by glutamate of astrocyte-neuron signalling.

As increased calcium concentration in astrocytes stimulates glutamate release, we used two methods that directly raise calcium in astrocytes and asked whether there was a consequent elevation in neuronal calcium. First, tapping astrocytes¹⁰ with micropipettes elevated astrocyte calcium levels and significantly raised neuronal calcium levels ($\Delta F/F_0 = 62 \pm 25\%$, $n = 6$; $P < 0.005$, Mann-Whitney *U*-test). Second, we photostimulated astrocytes to increase their calcium levels¹¹. A portion of a single astrocyte was exposed to focal ultraviolet light. Photostimulation reliably elevated astrocyte calcium levels ($\Delta F/F_0 = 202 \pm 19\%$, $n = 7$), which then spread to adjacent astrocytes (Fig. 3). When astrocytes were photostimulated, calcium also increased in neurons that were in contact with astrocytes

We used confocal microscopy to establish whether bradykinin causes a neuronal change in calcium level and to study the pharmacology of the receptor further. Co-cultures of astrocytes and neurons were loaded with Fluo-3 and the fluorescent emission at the plane of neuronal somata was monitored. Bradykinin reliably elevated neuronal calcium in astrocyte-neuron co-cultures (Fig. 2g). Addition of the NMDA receptor antagonist D-2-amino-5-phosphonopentanoic acid (D-AP5; 50 μ M), but

FIG. 4 Photostimulation of astrocytes causes an increase of $[Ca^{2+}]_i$ in adjacent unstimulated neurons owing to the action of glutamate. Images represent calcium levels in the confocal plane of the neuronal body before (a) and after (b) UV stimulation of a single astrocyte. After stimulation, $[Ca^{2+}]_i$ increased both in astrocytes (A) and in the neuron (N). The increase in neuronal calcium is not due to a direct action of UV exposure, because neuronal processes (arrows in c) were not in the path of UV excitation (circled lightning bolt). c, Microinjection of sulphorhodamine 101 (10 mM) into individual neurons ($n = 8$) confirmed that Fluo-3 revealed the full extent of neuronal processes. In control experiments ($n = 6$), we gradually moved the photostimulus closer to an isolated cell and determined that calcium levels were only elevated when the stimulus was within less than 5 μ m from the cell border. Neurites were always greater than 11 μ m from the edge of the photostimulus and thus were not directly activated by this stimulus. The image in c was acquired at the end of the experiment at the level of astrocytes and neuronal processes. The time course of $[Ca^{2+}]_i$ changes in neurons due to photostimulation (PS) of astrocytes is shown in d. The presence of DGG (1 mM), but not tetrodotoxin (TTX, 1 μ M), greatly reduced the neuronal response. Colour scale for a and b indicates linear pseudocolour representation of fluorescence intensity ranging from 0 to 160 units. Points represent means $\pm$ s.e.m. ($n = 15$ for each group). Scale bar, 10 μ m. METHODS. TTX (1 μ M) and DGG (1 mM) were uniformly applied to the bath by saline exchange 10–20 min before UV stimulation.



(Fig. 4). Photostimulation of astrocytes caused a maximal increase in neuronal $\Delta F/F_0$ ($81 \pm 20\%$; $n=15$) which was not significantly altered by addition of tetrodotoxin ($1 \mu\text{M}$; $\Delta F/F_0 = 57 \pm 6\%$, $n=15$; $P>0.25$, Scheffé's test) to the medium. However, addition of the glutamate receptor antagonist DGG significantly reduced the astrocyte-induced elevation of calcium in adjacent neurons ($\Delta F/F_0 = 14 \pm 4\%$, $n=15$; $P<0.01$, Scheffé's test; Fig. 4).

These results show that a neuropeptide, bradykinin, causes the calcium-dependent release of the neurotransmitter glutamate from neocortical astrocytes. The calcium-dependent glutamate release mechanism of astrocytes might be different from neuronal release mechanisms, because we found that astrocytes are not immunoreactive for the calcium-sensitive vesicle-protein synaptotagmin. Using three stimuli that each raise astrocyte calcium levels (bradykinin, tapping astrocytes with micropipettes, and direct astrocyte photostimulation), we have demonstrated that astrocytes can cause a glutamate-dependent elevation of neuronal calcium. These data point to a new form of transmitter signalling in the nervous system in which stimuli that mobilize calcium in astrocytes also cause a calcium-dependent release of the excitatory amino acid glutamate, which can then act on adjacent neurons. It is possible that other pathways, such as gap junctions¹², also contribute to astrocyte-neuron signalling. But as gap-junction uncoupling agents such as octanol have many other effects¹³⁻¹⁵, which include a reduction of calcium influx through the NMDA receptor¹³, we have not tested this possibility. Glutamate released at synaptic terminals can trigger calcium waves in an astrocyte network¹. As many neurotransmitters can mobilize glial calcium^{10,16-22}, it is possible that neurons stimulate neighbouring astrocytes to release glutamate, which in turn signals back to the neuron. In this case astrocytes may be an integral computational element within the brain. In addition to mediating fast synaptic transmission, glutamate serves a role in synaptic plasticity²³, and under conditions of excessive release can cause neuronal damage and degeneration²⁴. Astrocyte-mediated glutamate release could therefore be important in regulating neuronal physiology and pathology. □

Received 18 March; accepted 20 May 1994.

1. Dani, J. W., Chernjavsky, A. & Smith, S. J. *Neuron* **8**, 429-440 (1992).
2. Murphy, T. H., Blatter, L. A., Wier, W. G. & Baraban, J. M. *J. Neurosci.* **13**, 2672-2679 (1993).
3. Mennerick, S. & Zorumski, C. F. *Nature* **368**, 59-62 (1994).
4. Balcar, V. J. & Johnston, G. A. R. *J. Neurochem.* **19**, 2657-2666 (1972).
5. Isaacson, J. S. & Nicoll, R. A. *J. Neurophysiol.* **70**, 2187-2191 (1993).
6. Sokal, R. R. & Rohlf, F. J. in *Biometry* (ed. Wilson, J.) 400-453 (Freeman, New York, 1981).
7. Kimeberg, H. K., Goderie, S. K., Higman, S., Pang, S. & Waniewski, R. A. *J. Neurosci.* **10**, 1583-1591 (1990).
8. Jones, A. W., Smith, D. A. S. & Watkins, J. C. *Neuroscience* **13**, 573-581 (1984).
9. Scheffé, H. *Biometrika* **40**, 87-104 (1953).
10. Charles, A. C., Merrill, J. E., Dirksen, E. R. & Sanderson, M. J. *Neuron* **6**, 983-992 (1991).
11. Guthrie, P. B., Segal, M. & Kater, S. B. *Nature* **354**, 76-80 (1991).
12. Nedergaard, M. *Science* **263**, 1768-1771 (1994).
13. McLarnon, J. G., Wong, J. H. & Sawyer, D. Can. *J. Physiol. Pharmacol.* **69**, 1422-1427 (1991).
14. Triggie, D. J. et al. *Ann. N.Y. Acad. Sci.* **635**, 123-138 (1991).
15. Ma, J., Fill, M., Knudson, M. C., Campbell, K. P. & Coronado, R. *Science* **242**, 99-102 (1988).
16. Reiser, G., Donié, F. & Binmöller, F.-J. *J. Cell Sci.* **93**, 545-555 (1989).
17. Cornell-Bell, A. H., Finkbeiner, S. M., Cooper, M. S. & Smith, S. J. *Science* **247**, 470-473 (1990).
18. Glaum, S. R., Holzwarth, J. A. & Miller, R. J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3454-3458 (1990).
19. Cornell-Bell, A. H. & Finkbeiner, S. M. *Cell Calcium* **12**, 185-204 (1991).
20. Inagaki, N., Fukui, H., Ito, S. & Wada, H. *Neurosci. Lett.* **128**, 257-260 (1991).
21. Jensen, A. M. & Chiu, S. Y. *J. Neurosci.* **11**, 1674-1684 (1991).
22. McCarthy, K. D. & Salim, A. K. *Neuroscience* **41**, 325-333 (1991).
23. Malgaroli, A. & Tsien, R. W. *Nature* **357**, 134-139 (1992).
24. Choi, D. W. *Neuron* **1**, 623-634 (1988).
25. Levison, S. W. & McCarthy, K. D. in *Culturing Nerve Cells* (eds Banker, G. & Goslin, K.) 309-336 (MIT Press, Cambridge, MA, 1991).
26. Heutner, J. E. & Baughman, R. W. *J. Neurosci.* **6**, 3044-3060 (1986).
27. Parpura, V., Haydon, P. G. & Henderson, E. J. *Cell Sci.* **104**, 427-433 (1993).
28. Basarsky, T. A., Parpura, V. & Haydon, P. G. *J. Neurosci.* (in the press).
29. Kao, J. P. Y., Hartoontuan, A. T. & Tsien, R. Y. *J. Biol. Chem.* **263**, 8179-8184 (1988).

ACKNOWLEDGEMENTS. We thank D. S. Sakaguchi and S. Shen for comments on the manuscript; R. Jahn for antibody against synaptotagmin; D. S. Sakaguchi for help with immunocytochemistry; and S. Crabb and R. Loudon for artwork. This work was supported by the Epilepsy Foundation of America and NIH grants to P.G.H. and S.J., an MRC of Canada studentship to T.A.B. and by an NSF Research Training Group training grant. P.G.H. is a member of the ISU Laboratory of Cellular Signaling.

Myelination in the absence of myelin-associated glycoprotein

Chumei Li*, Michael B. Tropak*, Robert Gerlai*, Susan Clapoff*, Wanda Abramow-Newerly*, Bruce Trapp†, Alan Peterson‡ & John Roder*

* Samuel Lunenfeld Research Institute, 600 University Avenue, and The Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5G 1X5, Canada

† Department of Neurosciences, Cleveland Clinic Foundation, Cleveland, Ohio 44195, USA

‡ Royal Victoria Hospital, McGill University, 687 Pine Avenue, Montreal, Quebec H3A 1A1, Canada

THE hypothesis that myelin-associated glycoprotein (MAG) initiates myelin formation is based in part on observations that MAG has an adhesive role in interactions between oligodendrocytes and neurons¹. Furthermore, the over- or underexpression of MAG in transfected Schwann cells *in vitro* leads to accelerated myelination² or hypomyelination³, respectively. Here we test this idea by creating a null mutation in the *mag* locus and deriving mice that are totally deficient in MAG expression at the RNA and protein level. In adult mutant animals the degree of myelination and its compaction are normal, whereas the organization of the periaxonal

TABLE 1 Morphometric analysis of myelin

X-ray diffraction			
Tissue	MAG genotype		P
	-/-	+/+	
Sciatic nerve (nm)	17.4 ± 0.01 ²	17.4 ± 0.04	NS
Optic nerve (nm)	15.3 ± 0.12	15.2 ± 0.06	NS
Electron microscopy			
Optic nerve	MAG genotype		P
	-/-	+/+	
Total fibres	531	473	
Loss of collar (frequency)	453 (85.3%)	53 (11.2%)	<0.0001*
Periaxonal swelling (frequency)	90 (16.9%)	29 (6.1%)	<0.0001*
Cytoplasm within myelin (frequency)	70 (13%)	38 (8%)	<0.01*
Redundant myelin (frequency)	32 (6.0%)	6 (1.3%)	<0.0001*
Total fibres	183	173	
Myelin/axon	1.36 ± 0.08	1.04 ± 0.08	<0.0001†
Periaxonal swelling (area)	225 ± 27	52 ± 21	<0.0001†

X-ray diffraction: Freshly dissected nerves were sealed between mica windows 1 mm apart in phosphate-buffered saline. A bent quartz crystal monochromator was used to isolate the Cu(K α) line at $\lambda=1.54$ Å and diffraction of X-rays was measured at 25 °C in a Guinier camera⁴. Values represent the mean and s.e. of determinations of periodicity on intact nerves from several age-matched (3 months) and sex-matched littermates per group. P values were determined by Student's t-test. Electron microscopy: Optic nerves from a representative pair of perfused, age-matched (3 months) and sex-matched littermates were prepared for electron microscopy (Fig. 2 legend). Electron micrographs at $\times 15,504$ magnification were placed in montages, representing 1% of the optic nerve, digitized and analysed on a metamorph visual-imaging system. The frequency of inner cytoplasmic collars was scored for collars that surrounded more than half the axon diameter. Periaxonal swellings include normal inner lips (that is, mesaxon) and abnormal intrusions of glial cytoplasm. Cytoplasm within myelin-included Schmidt-Lanterman incisures and glial cytoplasmic intrusions that partially or completely surrounded the fibre diameter. Redundant myelin was manifested as extra loops of compact myelin sheaths that coursed away from the axons. A random subpopulation of fibres was scored for area ratios using the mean pixel area of myelin divided by axon area. Periaxonal areas were measured from the innermost lamella of compact myelin to the axon membrane. Values represent mean pixels $\pm$ s.e. from a digitized image.

* χ^2 analysis.

† Mann-Whitney U-test. Similar results were obtained blind by three independent observers on these and other paired +/+ and -/- mice.

FIG. 1 Disruption of the *mag* gene. **A**, The MAG targeting vector consisted of 7.5 kb of homology with a neomycin-resistance gene (*neo^R*) inserted into the *Xho* site (X) of exon V, and thymidine kinase gene (TK)²² into the *EcoRI* site (E) in exon VII. Wild-type alleles cut with *Bam*HI (B) and *Hind*III (H) generated two bands of 7.7 and 4.5 kb when probed with a 1.6-kb cDNA *Xho*I fragment encoding exons 5–10 (shaded). In the targeted allele the *neo^R* insertion generated a new *Bam*HI site and a diagnostic 5.5 kb fragment. **B**, Southern blot analysis on tail DNA from 134 offspring of heterozygous crosses. DNA from 33 wild-type mice showed the expected 7.7- and 4.5-kb bands, whereas 58 heterozygotes revealed a 5.5-kb diagnostic band and 43 homozygotes showed loss of the 7.7-kb wild-type band, following a mendelian distribution. RNA (≥ 0.5 kb) from brain (lower panel) was probed with MAG cDNA and a control northern blot revealed a normal MBP signal in the $-/-$ mice (not shown). **C**, Myelin tracts in mice perfused with 4% paraformaldehyde. One hemisphere from a $+/+$ and one from a $-/-$ brain were glued together, simultaneously sectioned and immunostained. Alternate coronal sections through the cortex were stained with antibody against MAG (a) or MBP (b). CC, corpus callosum. **D**, Western analysis. Proteins purified from myelin, crude brain extracts (CNS), sciatic or trigeminal nerves (PNS) were probed with the indicated antibodies. A difference of less than 1.5- to 2-fold would not be detected. MAG $-/-$ mice were paired with age (3 months) matched MAG $+/+$ or MAG $+/-$ littermate controls in all experiments, except the L1 PNS western mice, which were 3 weeks old.

METHODS. Isogenic (129) vectors, yielded 5/511 targeted events²³ in our ES cell line, R1 (ref. 24), but non-isogenic vectors gave 0/540, thereby confirming the importance of isogenic DNA^{25,26}. Chimaeric mice were made from one line of ES cells following aggregation with CD1 morulae²⁷.

region is partially impaired. Mutant animals show a subtle intention tremor. Our findings do not support the widely held view that MAG is critical for myelin formation but rather indicate that MAG is necessary for maintenance of the cytoplasmic collar and periaxonal space of myelinated fibres.

The targeted disruption of the *mag* gene yielded homozygous null mutant mice ($-/-$) which showed a complete absence of MAG messenger RNA and protein in the central nervous system (CNS), whereas myelin basic protein and proteolipid protein were present in normal amounts (Fig. 1A, B, D). Nerves from the peripheral nervous system (PNS) of $-/-$ mice had also lost MAG, whereas protein O was not affected. MBP was selectively upregulated 1.5-fold in the PNS. Other proteins, including L1 and the neural cell-adhesion molecule NCAM, were normal in both the PNS and CNS of MAG mutants. Myelin tracts from $-/-$ mice formed normally in the absence of MAG, as demonstrated by immunocytochemistry (Fig. 1C).

The lamellar organization of the myelin in the $-/-$ mice was investigated using X-ray diffraction⁴ (Table 1). The mean periodicity was 17.4 nm for sciatic nerve and 15.3 nm for optic nerve, with no significant difference between the $+/+$ and $-/-$ littermates. As a large population of axons was averaged by this technique, this suggests that myelin compaction in the MAG mutant mice was normal. There were no striking differences evident in light microscopy in the myelination of any CNS or PNS fibres in $+/+$ or $-/-$ mice (Fig. 2a, b). In electron micro-

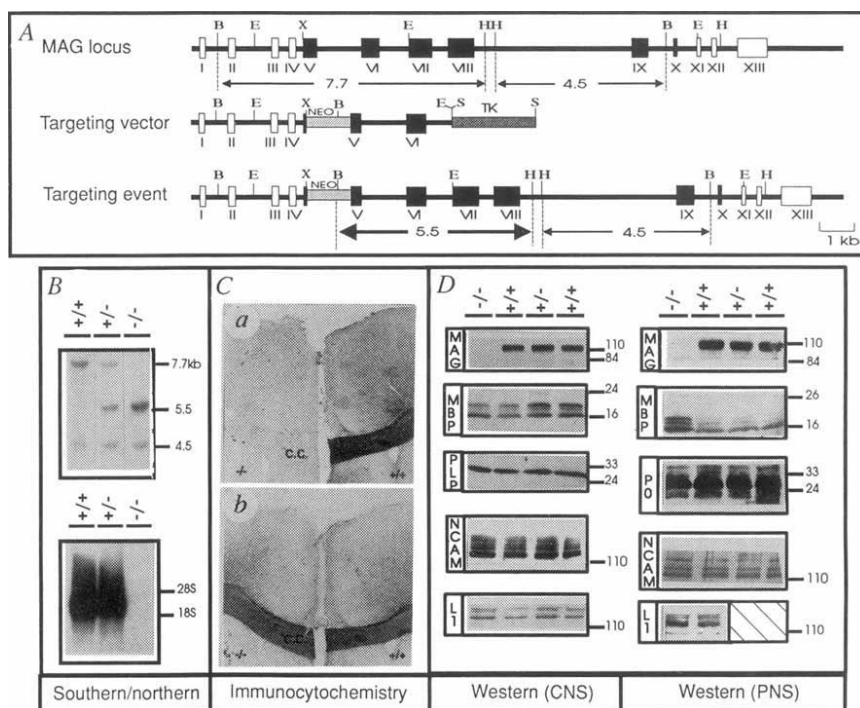


TABLE 2 Motor and posture patterns

Open field			
Behaviour	Mutant (−/−)	Control (+/+, +/−)	<i>P</i>
Locomotion score (frequency)	160.5 ± 25.24	181.4 ± 30.94	>0.5
Grooming (% time)	6.44 ± 1.63	5.89 ± 1.04	>0.5
Leaning against the wall (frequency)	11.74 ± 2.98	17.17 ± 3.90	>0.1
Bar cross			
Behaviour	Mutant (−/−)	Control (+/+, +/−)	<i>P</i>
Crossing challenge bar (% time)	20.6 ± 3.41	10.1 ± 1.61	<0.05
Locomotion on wide bar (% time)	11.4 ± 1.61	18.4 ± 2.44	<0.05
Grooming (% time)	1.32 ± 0.84	2.81 ± 0.98	<0.05
Sniffing down (frequency)	12.44 ± 1.60	19.11 ± 1.85	<0.05
Intention tremor (% time)	2.67 ± 1.02	0.00 ± 0.00	<0.0001
Intention tremor (frequency)	12 out of 18	0 out of 18	<0.0001

Spontaneous motor and posture patterns were recorded in two novel situations: open field and a bar-cross apparatus⁸. Age- and sex-matched mice (8 weeks) were tested singly and in randomized order (mutant, $n=19$; control, $n=18$). The open field was a plastic box (46 $\times$ 25 $\times$ 15 cm) with its bottom covered with a sheet of paper with a 5-cm² grid pattern allowing the recording of locomotion scores. In this test, exploratory activities, motor and posture patterns were monitored for 10 min. 'Locomotion score', number of squares crossed on the floor grid; a cross was counted when the mouse entered a new square with both forepaws. 'Grooming', stereotypical movements, including face cleaning with the forepaws and fur licking. 'Leaning', leaning against the wall of the open field box with one or both forepaws while standing on the hindlegs. The bar-cross apparatus was an elevated U-shaped platform with two wider (18 mm) bars and a connecting narrow (1 mm wide) challenge bar. In this situation, finer motor coordination and balance were tested for 3 min by recording motor posture patterns. 'Locomotion on the wide bars', locomotion activity; 'sniffing down', holding the head below the level of the platform with the nose pointing downwards with typical whisker movements; 'intention tremor', short (1–3 s) transient tremor of the trunk immediately preceding active movement. Mice were then placed directly onto the challenge bar as they would not cross it voluntarily. The per cent of time taken to complete the crossing was measured ('crossing the challenge bar'). Values represent the mean $\pm$ s.e. per cent of time or frequencies of behaviours. Statistical significance was calculated by a non-parametric Mann-Whitney test for all values except the last, where a Fisher's exact test was used.

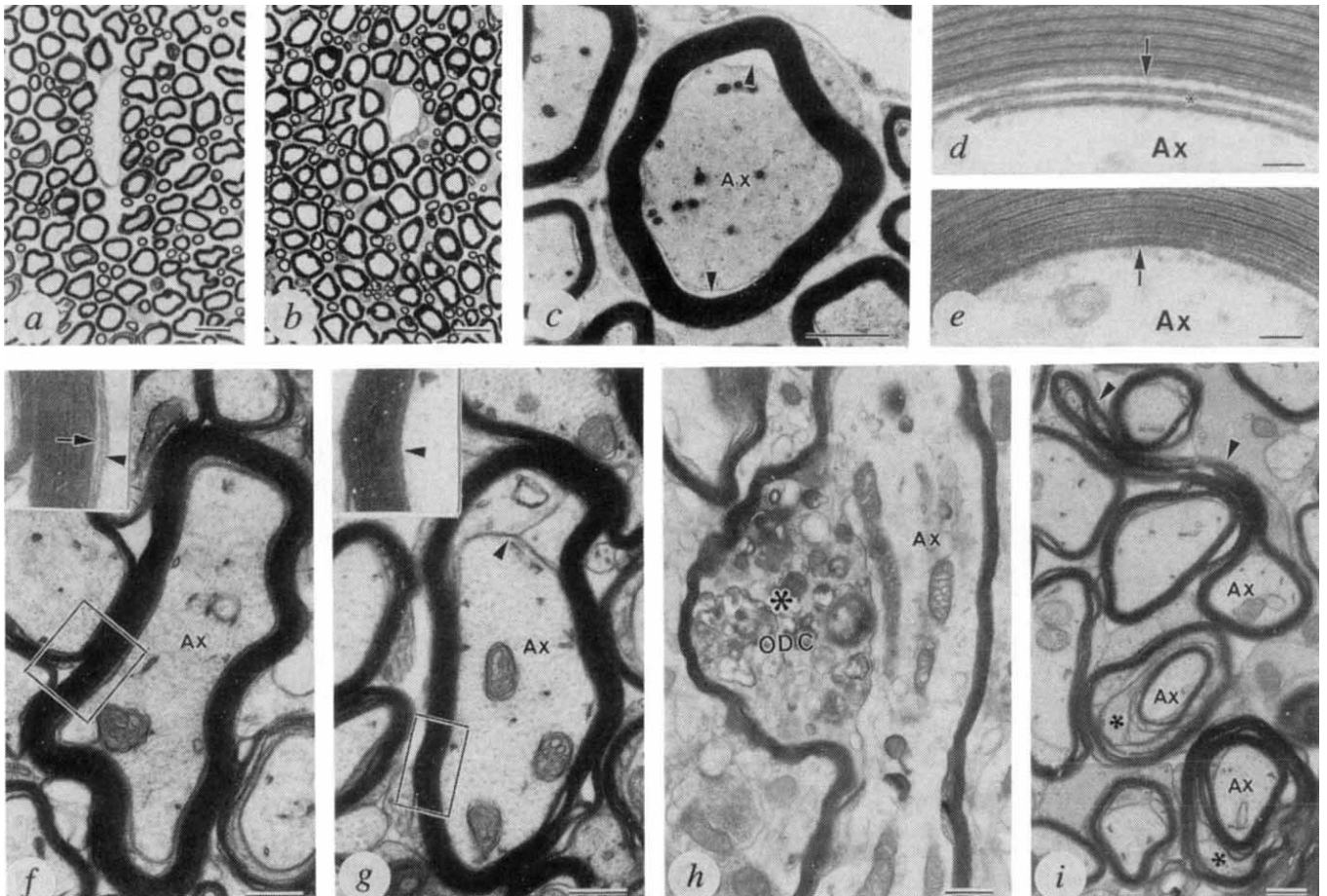


FIG. 2 Myelin morphology in PNS (upper panels) and CNS (lower panels). *a* and *b*, light micrographs of 1.0- μ m sections of L2 ventral spinal roots of $MAG^{+/+}$ (*a*) and $MAG^{-/-}$ (*b*) mice. There were no striking differences in the number of myelinated fibres or in the thickness of the myelin sheaths or in the calibre of the axons between these two groups. *c*, Electron micrograph of a myelinated fibre from the L4 ventral root of a $MAG^{-/-}$ mouse displaying a dilated periaxonal space (arrowheads). *d* and *e*, Electron micrographs of the normal 12–14-nm periaxonal space (*d*, asterisk) and the Schwann cell cytoplasmic collar (*d*, arrow) from the L2 ventral root of a $MAG^{+/+}$ mouse, features not preserved in the $MAG^{-/-}$ sample (*e*, arrow). The Schmidt–Lanterman incisures, paranodal regions and outer mesaxons were present in the MAG mutant mice (data not shown). *f* and *g*, Electron micrographs of $MAG^{+/+}$ optic nerve fibre (*f*) with normal periaxonal space (insert, arrowhead) and cytoplasmic collar (insert, arrow) and a $MAG^{-/-}$ optic nerve fibre (*g*) with

decreased periaxonal space (inset, arrowhead), and the absence of the oligodendrocyte cytoplasmic collar (inset, arrowhead) except for the mesaxon region (arrowhead). *h*, Longitudinal section of a $MAG^{-/-}$ optic nerve fibre with focal disorganization of the periaxonal cytoplasmic collar (asterisk). *i*, Transverse section of $MAG^{-/-}$ optic nerve fibres showing disrupted compact myelin lamella (asterisks) and redundant compact myelin (arrowheads). Age- (3 months) and sex-matched control littermates were used in all experiments. Ax, axon; ODC, oligodendrocyte. Scale bars and magnification: *a*, *b*, 10 μ m; *c*, 0.5 μ m; *d*, *e*, 0.05 μ m; *f*, *g*, 0.5 μ m, insets twice as large; *h*, 0.5 μ m; *i*, 0.5 μ m.

METHODS. Animals were perfused with 0.5% paraformaldehyde and 2.5% glutaraldehyde in PBS, post-fixed in OsO_4 , embedded in Epon, sectioned (<1 μ m) and stained with lead acetate and uranyl acetate.

graphs at higher magnification (Fig. 2*c*), however, some axon profiles were separated from the inner aspect of their myelin sheaths. Although fixation artefacts could account for apparent abnormalities in some fibres, the 12–14 nm periaxonal space is very resilient and does not come apart as a result of gross axonal swelling after chemical intoxication⁵ and is also stable in the face of the severe axonal shrinkage that occurs in mutant Syrian hamsters⁶. Examples of gross perturbations in morphology were often seen in MAG mutants, but no single abnormality was consistently observed in every fibre. In the MAG mutants, the periaxonal cytoplasmic collar was frequently missing in optic nerve (Fig. 2*f*, *g*), cervical spinal cord (not shown) and ventral roots (Fig. 2*d*, *e*), and the cytoplasmic leaflet of the periaxonal membrane had often fused with the inner compact myelin lamellae to form a major dense line (Fig. 2*d*, *e*). When present, the cytoplasm of the periaxonal collar was usually disorganized (Fig.

2*h*). Profiles of redundant compact myelin coursing away from the axon (Fig. 2*i*) were seen more frequently in the mutant (Table 1), and as such redundant myelin is occasionally found in the development of normal fibres, this might reflect a retardation in myelin maturation. Owing to redundant myelin and swollen collars, the ratio of myelin area to axon area was larger in the mutant than in the wild type. Significantly more ($P < 0.01$) optic nerve fibres in $-/-$ mice, compared to $+/+$ mice, had lost cytoplasmic collars, swollen collars or redundant myelin (Table 1).

The motor and posture patterns⁷ of the mutant mice were the same as those of the normal mice in the open field, but differences were noted in the bar-cross apparatus⁸, which requires finer motor coordination (Table 2). Mutant mice showed decreased locomotor activity, they were slower to traverse the narrow challenge bar, and exploratory sniffing and grooming were less fre-

quent, behaviours that require good balance on the bar-cross apparatus. Furthermore, a significant number of mutant mice exhibited a mild, transient, trunk tremor on the bar that was never observed in the normal mice ($P < 0.0001$). These results suggest that although no gross behavioural abnormalities were apparent in the MAG mutant mice, their finer motor coordination abilities were significantly affected by the mutation.

Our results indicate that MAG is not essential for myelin formation as the number of lamellae and the periodicity of the compact myelin appeared normal in MAG-deficient mutants. Rather, MAG may be involved in the organization of the periaxonal space and periaxonal cytoplasmic collar. The enrichment of MAG in the periaxonal membrane of all myelin internodes^{9,10}, its exclusion from compact myelin, and the strict correlation between the presence of MAG and the maintenance of the periaxonal space and periaxonal cytoplasm in quaking mutants, all support this hypothesis¹¹. The absence of a more profound phenotype could arise by overexpression of compensating molecules in the MAG mutants. For example, L1 and NCAM are expressed early on axons and Schwann cells, whereas MAG expression initiates as L1 and NCAM are shut off^{10,12}. It was conceivable, therefore, that loss of MAG would induce the continued expression of L1 or NCAM. However, there was no increase (>2-fold) in either protein in $-/-$ mice (Fig. 1) and none was detected in adult paranodes or incisures (data not shown).

Several observations suggest that MAG may act as a glue or spacer for the periaxonal region¹³. First, MAG belongs to the immunoglobulin superfamily¹⁴⁻¹⁷, sharing homology with NCAM and CD22, which are known to mediate cell-cell adhesion. Second, MAG can bind axons when incorporated into liposomes^{1,18}. Third, loss of MAG leads to the collapse of the periaxonal collar. It is conceivable that the cytoplasmic portion of MAG generates a signal^{19,20} to maintain this collar. Fyn may mediate the signal since Fyn phosphorylates MAG and *fyn*⁻ mice were hypomyelinated²¹. The breeding of MAG-deficient mice with other mutants will help distinguish between these possibilities. □

Specificity and flexibility in thymic selection

Stephen C. Jameson, Kristin A. Hogquist
& Michael J. Bevan

Howard Hughes Medical Institute, Department of Immunology,
University of Washington, Seattle, Washington 98195, USA

DURING positive selection, developing thymocytes are rescued from programmed cell death by T-cell receptor (TCR)-mediated recognition of major histocompatibility complex (MHC) molecules¹⁻³. MHC-bound peptides contribute to this process⁴⁻⁸. Recently we identified individual MHC-binding peptides which can induce positive selection of a single TCR⁹. Here we examine peptide fine specificity in positive selection. These data suggest that a direct TCR-peptide interaction occurs during this event, and strengthens the correlation between selecting peptides and TCR antagonists^{9,10}. Certain positively selecting peptides are weakly antigenic⁹. We demonstrate that thymocytes 'educated' on such a peptide are specifically non-responsive to it and have decreased CD8 expression levels. Similar reduction of CD8 expression on mature T cells converts a TCR agonist into a TCR antagonist. These data indicate that thymocytes may maintain self-tolerance towards a positively selecting ligand by regulating co-receptor expression.

Positive selection of T cells occurs at the CD4⁺8⁺ stage of development and results in phenotypic changes (including down-

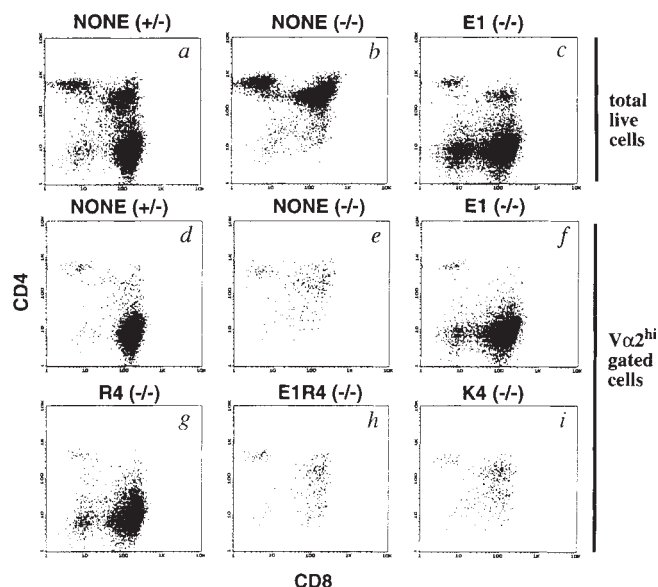


FIG. 1 Flow cytometric analysis of thymocytes from fetal thymic organ cultures of H-2^b OVA-tcr-1 mice. Thymi from $\beta_2M^{+/-}$ (a, d) or $\beta_2M^{-/-}$ fetuses (b, c, e-i) were cultured with the indicated peptides. Thymocytes were stained for expression of CD4, CD8 and the transgenic V α 2 region (V α 2^{hi}). CD4 versus CD8 staining is shown for all live gated cells (a-c) or V α 2^{hi} cells (d-i).

METHODS. FTOC conditions were essentially as described⁹. Thymus lobes were recovered from fetuses at day 16 of gestation, and cultured in RP10 media containing human β_2M (CalBiochem) (5 μ g ml⁻¹) and the indicated peptides (at 20 μ M), with daily replenishment for 7 days. Peptide names refer to the amino-acid substitutions in OVAp (sequence SIINFEKL). Thus, E1 has a Ser→Glu substitution at position 1. All peptides were synthesized on a Synergy peptide synthesizer (Applied Biosystems). $\beta_2M^{+/-}$ and $-/-$ lobes were distinguished by staining with Y3 (anti-K^b). Anti-CD4-PE and anti-CD8-FITC were obtained from Becton Dickinson and used directly. The anti-V α 2 antibody B20 (ref. 25) was conjugated to biotin and staining revealed with Tri-colour-streptavidin (Caltag). Analysis was done on a FACScan (Becton-Dickinson) on 20,000 live gated cells.

Received 28 April; accepted 3 June 1994.

- Poltorak, M. et al. *J. Cell Biol.* **105**, 1893-1899 (1987).
- Owens, G. C., Boyd, C. J., Bunge, R. P. & Salzer, J. L. *J. Cell Biol.* **111**, 1171-1182 (1990).
- Owens, G. C. & Bunge, R. P. *Neuron* **7**, 565-575 (1991).
- Kirschner, D. A., Ganser, A. L. & Caspar, D. L. in *Myelin* 2nd edn (ed. Morell, P.) Ch. 2, 51-95 (Plenum, London, 1984).
- Trapp, B., Quarles, R. & Griffin, J. J. *J. Cell Biol.* **98**, 1272-1278 (1984).
- Hirano, A. *Prog. Neuropath.* **5**, 99-112 (1983).
- Huntingford, F. *The Study of Animal Behaviour* (Chapman and Hall, London, 1984).
- Gerlai, R. et al. *Behav. Brain Res.* **55**, 51-59 (1993).
- Sternberger, N. H., Quarles, R. H., Hoyama, Y. & Webster, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1510-1514 (1979).
- Martini, R. & Schachner, M. *J. Cell Biol.* **103**, 2439-2448 (1986).
- Trapp, B. D., Quarles, R. H. & Suzuki, K. J. *J. Cell Biol.* **99**, 594-606 (1984).
- Wood, P. M., Schachner, M. & Bunge, R. P. *J. Neurosci.* **10**, 3635-3645 (1990).
- Trapp, B. D. & Quarles, R. H. *J. Cell Biol.* **92**, 877-882 (1982).
- Aquino, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 600-604 (1987).
- Lai, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4337-4341 (1987).
- Salzer, J. L., Holmes, W. P. & Colman, D. R. *J. Cell Biol.* **104**, 957-965 (1987).
- Dulac, C. et al. *Neuron* **8**, 1-20 (1992).
- Johnson, P. W. et al. *Neuron* **3**, 377-385 (1989).
- Tropak, M. B., Johnson, P. W., Dunn, R. J. & Roder, J. C. *Molec. Brain Res.* **4**, 143-155 (1988).
- Edwards, A. M. et al. *Molec. cell. Biol.* **8**, 2655-2658 (1988).
- Uemori, H. et al. *Nature* **367**, 572-576 (1994).
- Soriano, P., Montgomery, C., Geske, R. & Bradley, A. *Cell* **64**, 693-702 (1991).
- Capecchi, M. *Sci. Am.* **270**, 52-59 (March, 1994).
- Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8424-8428 (1993).
- Te Riele, H., Robanus, E. & Berns, A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5128-5132 (1992).
- Van Deursen, J. & Wieringa, B. *Nucleic Acids Res.* **20**, 3815-3820 (1992).
- Wood, S. A., Allen, N. D., Rossant, J., Auerbach, A. & Nagy, A. *Nature* **365**, 87-89 (1993).

ACKNOWLEDGEMENTS. We thank P. Rand, N. Fuller, N. Seniuk, I. Gitelman, S. Ludwin, A. Roach, W. Tatton, B. Esposito, P. Valera and C. Haney for assistance. This work was supported by MRC Canada, the Amalgamated Transit Union MS Research Fund, the NIH and The Canadian Networks of Centres of Excellence on Neuroscience. J.R. is an MRC scientist. R.G. is an MRC Fellow, and C.L. and M.T. were supported by MSS Studentships.

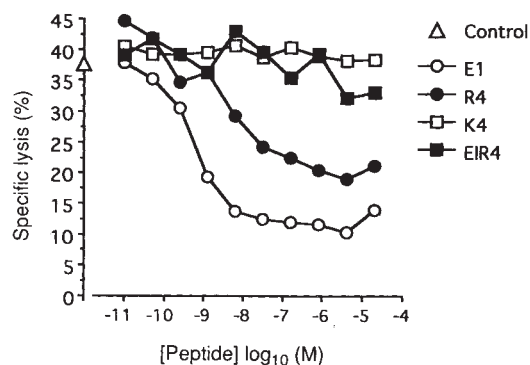


FIG. 2 TCR antagonist properties of OVAp analogues used in FTOC. The CTL line TG-0 is derived from an OVA-*tcrl* mouse and has been described before⁹. Lysis of EL4 target cells pre-pulsed with 2 pM OVAp was assayed in the presence or absence of the indicated peptide variants in a 4-h ⁵¹Cr-release assay. All four variant peptides bind K^b to a similar extent⁹ (data not shown). Specific lysis in the absence of variants was 37% and is shown on the abscissa.

regulation of CD4 or CD8) and functional maturation^{2,3,11,12}. We previously described a TCR-transgenic mouse (OVA-*tcrl*) with a receptor specific for ovalbumin residues 257–264 (OVAp) plus H-2K^b. Thymocytes bearing this receptor are positively selected in fetal thymic organ cultures (FTOCs) expressing K^b (Fig. 1a, d) (ref. 9) but maturation is impaired in FTOCs from $\beta_2M^{-/-}$ mice (Fig. 1b, e), where class I expression is drastically reduced^{7,13,15}. Selection is restored by addition of exogenous β_2 -microglobulin (β_2M) and certain variants of OVAp, exemplified by peptides E1 (Fig. 1c, f) and R4 (Fig. 1g) (ref. 9). This process shows exquisite peptide fine specificity; neither E1R4 (which combines both the E1 and R4 substitutions) nor K4 (which is chemically very similar to R4) are capable of positively selecting the OVA-*tcrl* receptor (Fig. 1h, i). Two theories have been pro-

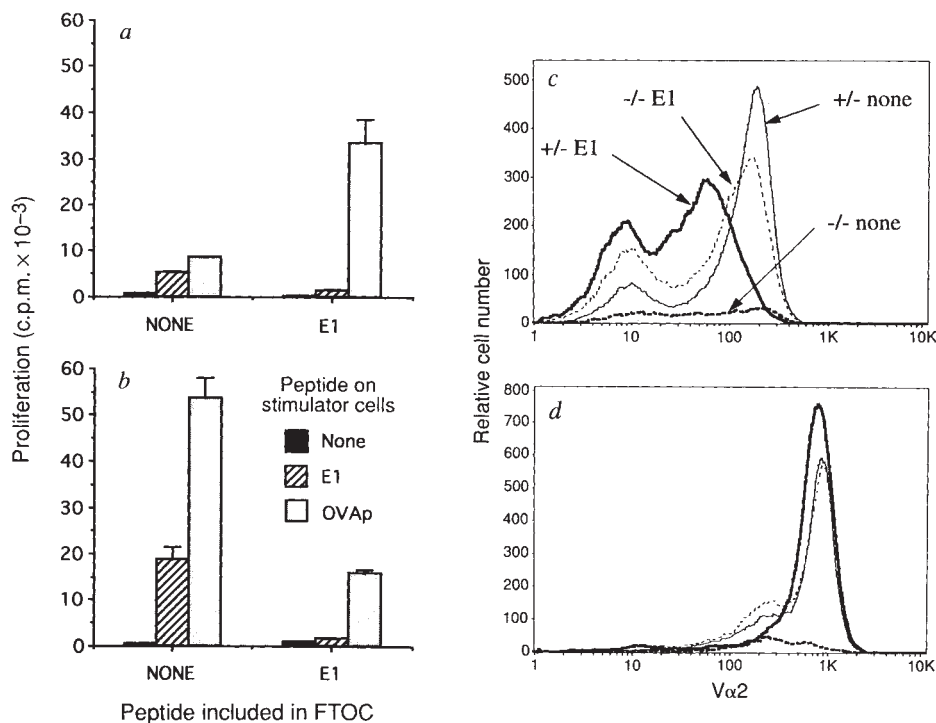
posed to explain peptide specificity in positive selection^{8,9,16}. The TCR may interact directly with both MHC and peptide residues ('peptide interaction' model), or it may only recognize MHC residues, in which case the peptide's role is merely to avoid impeding the TCR-MHC interaction ('peptide avoidance' model). Our results are hard to reconcile with the latter model; thus if both E1 and R4 simply keep out of the TCR's way, it is likely that E1R4 would do so too, and hence should positively select. Therefore, our data support a peptide interaction model.

We previously showed that peptides that induce positive selection, including E1 and R4, also act as TCR antagonists^{9,10} (Fig. 2). K4 and E1R4 did not behave as antagonists (Fig. 2), demonstrating similar fine specificity for positive selection and TCR antagonism. This suggests that the two phenomena may have a similar mechanistic basis, possibly based on TCR affinity for the peptide-MHC ligand^{9,10,17,18}. However, some peptides that are weak agonists induce positive selection at low concentrations and negative selection at higher concentrations (refs 9, 19, 20 and K.A.H. *et al.*, manuscript in preparation). Thus, positive selection is conditioned by both qualitative (perhaps affinity) and quantitative (ligand density) aspects of the TCR interaction.

E1 not only antagonizes OVA-*tcrl* T-cell responses (Fig. 2), but is also a weak agonist, stimulating mature transgenic cytotoxic T lymphocytes (CTLs)⁹ and deleting immature thymocytes in suspension culture (data not shown). We investigated how the presence of E1 during development alters the reactivity of mature OVA-*tcrl* thymocytes (Fig. 3a, b). Thymocytes from TCR-transgenic $\beta_2M^{-/-}$ FTOCs made a very weak response to both OVAp and E1. Addition of E1 during FTOC increased the absolute numbers of CD8⁺ T cells per lobe and thus the response to OVAp. But these cells were virtually non-responsive to E1 (Fig. 3a). This lack of E1 reactivity is all the more remarkable as the antigen-presenting cells used are $\beta_2M^{+/+}$ and so present peptides 50–100-fold better than $\beta_2M^{-/-}$ cells^{7,9}. Similarly, inclusion of E1 in OVA-*tcrl* $\beta_2M^{+/+}$ FTOC ablated the E1 response but only partially reduced proliferation to OVAp (Fig. 3b).

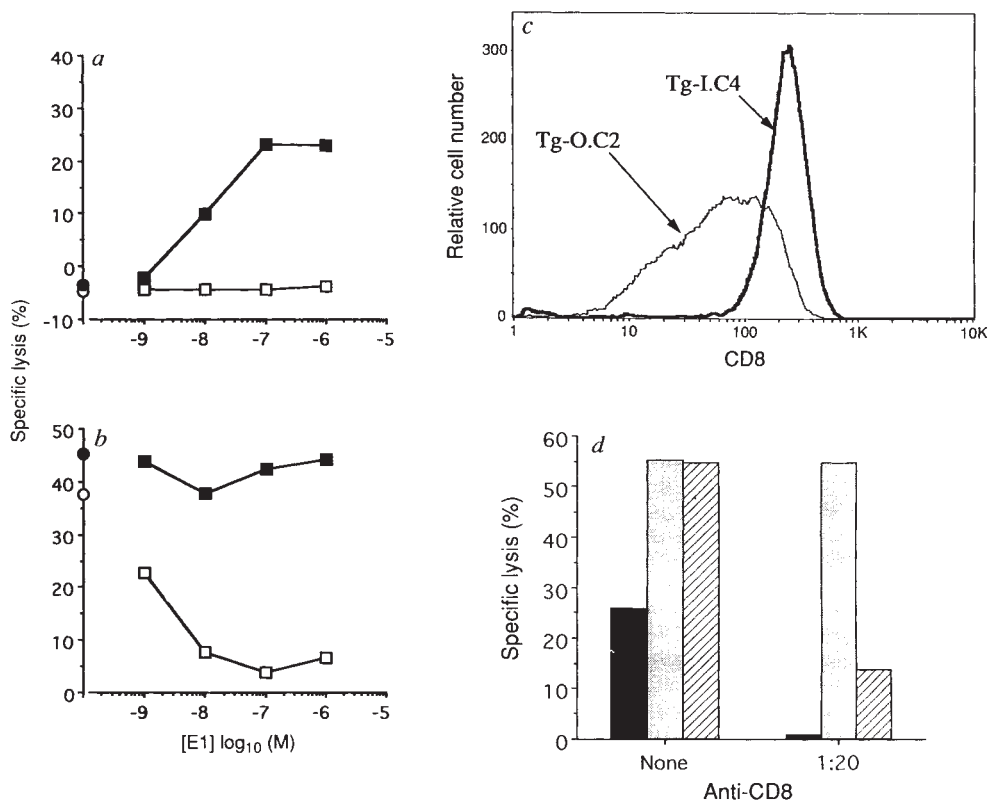
FIG. 3 Functional and phenotypic analysis of OVA-*tcrl* FTOCs cultured with and without peptide E1. $\beta_2M^{-/-}$ (a) and $\beta_2M^{+/+}$ (b) OVA-*tcrl* FTOCs were cultured with or without E1 at 20 μ M as indicated. Thymocytes were recovered (four lobes pooled in each group). Numbers ($\times 10^5$) of V α 2^{hi} CD4⁺ 8⁺ cells from each group were $\beta_2M^{+/+}$, 1.50; $\beta_2M^{+/+}$ plus E1, 0.75; $\beta_2M^{-/-}$, 0.16; and $\beta_2M^{-/-}$ plus E1, 1.19. a, b, The proliferative response of FTOC thymocytes to EL4 cells, with or without E1 (20 μ M) or OVAp (10 nM) pre-coating was determined by ³H-thymidine incorporation. Data show average of triplicate cultures with standard deviation. c, d, An aliquot of the cells recovered from each FTOC was analysed by FACS for expression of V α 2, CD4 and CD8; 10,000 live thymocytes were analysed. After gating for CD4⁺ cells, CD8 (c) and V α 2 (d) expression was determined. The same symbols are used in c and d. The average ($n = 4$) mean channel fluorescence for CD8 expression on CD8 single positive cells in this experiment was $\beta_2M^{+/+}$, 154 (± 8); $\beta_2M^{+/+}$ plus E1, 74 (± 6) ($P < 0.001$ versus $\beta_2M^{+/+}$) and $\beta_2M^{-/-}$ plus E1, 120 (± 23) ($P < 0.05$ versus $\beta_2M^{+/+}$). In an independent experiment the values were $\beta_2M^{+/+}$, 142 (± 3); $\beta_2M^{+/+}$ plus E1, 60 (± 6) ($P < 0.001$ versus $\beta_2M^{+/+}$) and $\beta_2M^{-/-}$ plus E1, 115 (± 9) ($P < 0.002$ versus $\beta_2M^{+/+}$).

METHODS. As for Fig. 1 except for the proliferation assay which was as in ref. 9. Briefly, irradiated EL4 stimulator cells were pulsed with peptide for 1 h and washed. One-twelfth of the pooled thymocytes in each group



was incubated with 10^4 stimulator cells in triplicate cultures. ³H-thymidine was added at 48 h and the culture collected 6 h later.

Fig. 4 Functional and phenotypic analysis of OVA-*tcr-1*-derived CTL clones. The clones TG-O.C2 and TG-I.C4 were obtained by limiting dilution from OVA-*tcr-1*-derived CTL lines. Reactivity of TG-O.C2 (open symbols) and TG-I.C4 (filled symbols) toward E1 was measured on *a*, EL4 (agonist assay) or *b*, EL4 prepulsed with 6 pM OVAp (antagonist assay) in a 4-h ^{51}Cr -release experiment. Symbols on the abscissa represent the lysis in the absence of added E1 peptide. *c*, CD8 expression levels on TG-O.C2 and TG-I.C4 (methods as in Fig. 1). *d*, Effect of anti-CD8 antibody on the specificity of TG-I.C4. Target cells were EL4 cells plus 1 μM E1 (solid bars), pre-pulsed with 6 pM OVAp (stippled bars) or pre-pulsed with 1 μM E1 (hatched bars). The IgG anti-CD8 antibody 53-6 was added (1/20 dilution of tissue culture supernatant) for the duration of the assay where indicated. TG-I.C4 was added at an effector:target ratio of 5:1 and lysis measured after 4 h. Inhibition of the responses by addition of anti-CD8 are: EL4-E1, 98%; EL4-OVAp, 1%; EL4-OVAp-E1, 75%.



What is the basis of this selective non-reactivity to E1? We investigated whether thymocyte TCR and/or CD8 expression levels were affected. TCR- α -chain expression was essentially equivalent in all the populations (Fig. 3d), but CD8 expression was significantly lowered on thymocytes from FTOCs in which E1 was included (Fig. 3c). This is seen most dramatically in the $\beta_2\text{M}^{+/-}$ FTOC, but is also apparent in cells from $\beta_2\text{M}^{-/-}$ lobes in which E1 is the positively selecting ligand. These data suggest that thymocytes can regulate co-receptor expression levels to avoid overt reactivity to the positively selecting ligand. Mature cells bearing low levels of CD8 (CD8^{lo}) could arise from selective survival of CD8^{lo} precursors or by CD8 downregulation on CD8^{hi} precursors. Relative to $\beta_2\text{M}^{+/-}$ FTOC we observed a decrease in the number of TCR $^{\text{hi}}$ $\text{CD8}^{\text{+}}$ cells in $\beta_2\text{M}^{-/-}$ plus E1 (~20%) or $\beta_2\text{M}^{+/-}$ plus E1 (~50%) cultures (see Fig. 3 legend), which supports the former model but does not exclude the latter.

To test whether CD8 expression levels alone could explain the differential response to E1 and OVAp, we analysed mature T-cell clones from OVA-*tcr-1* animals (Fig. 4a, b). TG-O.C2 is typical of most clones in that it responded to E1 as an antagonist. However, TG-I.C4 responded to E1 as a weak agonist. This response is completely blocked by an antibody to the transgenic α -chain (data not shown), suggesting that E1 reactivity is not due to expression of an endogenous TCR α -chain. Fluorescence-activated cell sorting (FACS) analysis indicated that CD8 (Fig. 4c) and TCR (data not shown) expression was higher on TG-I.C4 than TG-O.C2. To assess CD8 contribution in E1 reactivity, we tested the effects of anti-CD8 antibody. Partial CD8 blocking dramatically inhibited the TG-I.C4 response to E1 but, at this dose, did not inhibit OVAp reactivity (Fig. 4d). Interestingly, this treatment resulted in E1 now acting as a potent antagonist for TG-I.C4. Thus, the CD8 interaction can apparently determine whether a ligand is perceived as an agonist or an antagonist.

During positive selection, a T cell must interact with self MHC while avoiding negative selection on one hand and autoreactivity on the other. This could be achieved by regulation of receptor expression levels and/or signal transduction machinery. Our

data suggest that flexibility in CD8 expression levels is one way for thymocytes to fine-tune the efficacy of their interaction with MHC-peptide ligands present in the thymus and achieve positive selection while maintaining self tolerance. In what may be an analogous situation, lowering CD8 expression on mature T cells converts TCR agonists into antagonists. We note that selection by agonist ligands in a different TCR transgenic model yields CD8^{lo} single positive thymocytes also, although the functional potential of these cells was not addressed^{19,20}. In keeping with our observations, CD8 overexpression apparently converts a positively selecting environment into a negatively selecting one^{21,22}. Also, CD8 and/or TCR downregulation in peripheral T cells allows survival in the face of potentially tolerizing antigens^{23,24}. □

Received 11 February; accepted 6 May 1994.

- Bevan, M. J. *Nature* **269**, 417-418 (1977).
- Sprent, J. & Webb, S. R. *Adv. Immun.* **41**, 39-133 (1987).
- Rothenberg, E. V. *Adv. Immun.* **51**, 85-214 (1992).
- Nikolic Zugić, J. & Bevan, M. J. *Nature* **344**, 65-67 (1990).
- Sha, W. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6186-6190 (1990).
- Berg, L. J., Frank, G. D. & Davis, M. M. *Cell* **60**, 1043-1053 (1990).
- Hogquist, K. A., Gavin, M. A. & Bevan, M. J. *J. exp. Med.* **177**, 1469-1473 (1993).
- Ashton-Rickardt, P. G., Van Kaer, L., Schumacher, T. N. M., Ploegh, H. L. & Tonegawa, S. *Cell* **73**, 1041-1049 (1993).
- Hogquist, K. A. et al. *Cell* **76**, 17-28 (1993).
- Jameson, S. C., Carbone, F. R. & Bevan, M. J. *J. exp. Med.* **177**, 1541-1550 (1993).
- Blackman, M., Kappler, J. & Marrack, P. *Science* **248**, 1335-1341 (1990).
- von Boehmer, H. A. *Rev. Immun.* **8**, 531-556 (1990).
- Zijlstra, M. et al. *Nature* **344**, 742-746 (1990).
- Koller, B. H., Marrack, P., Kappler, J. W. & Smithies, O. *Science* **248**, 1227-1230 (1990).
- Vitiello, A., Potter, T. A. & Sherman, L. A. *Science* **250**, 1423-1426 (1990).
- Moller, G. (ed.) *Immun. Rev.* **135** (1993).
- Alexander, J. et al. *J. Immun.* **150**, 1-7 (1993).
- Evavold, B. D., Sloan-Lancaster, J. & Allen, P. M. *Immun. Today* **14**, 602-609 (1993).
- Ashton-Rickardt, P. G. et al. *Cell* **76**, 651-663 (1994).
- Sebzda, E. et al. *Science* **263**, 1615-1618 (1994).
- Lee, N. A., Loh, D. Y. & Lacy, E. J. *exp. Med.* **175**, 1013-1025 (1992).
- Robey, E. A. et al. *Cell* **69**, 1089-1096 (1992).
- Hammerling, G. et al. *Immun. Rev.* **122**, 47-67 (1991).
- Auphan, N. et al. *Int. Immun.* **4**, 1419-1428 (1992).
- Grégoire, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8077-8081 (1991).

ACKNOWLEDGEMENTS. We thank F. Carbone and B. Heath for supplying the TCR transgenic mice and conversation and M. Zollman and K. McConnell for technical assistance. This work was supported by NIAID and HHMI. S.C.J. is a Special Fellow of the Leukemia Society of America.

Negative feedback regulation of IgE synthesis by murine CD23

Philipp Yu*, Marie Kosco-Vilbois†, Mark Richards‡, Georges Köhler* & Marinus C. Lamers*§

* Max-Planck-Institut für Immunbiologie, Stübeweg 51, D-79108 Freiburg, Germany

† Basel Institute for Immunology, Grenzacherstrasse 487, 4005 Basel, Switzerland

‡ Medical Biology Institute, 11077 North Torrey Pines Road, La Jolla, California 92037, USA

§ To whom correspondence should be addressed

IMMUNOGLOBULIN E is found in nanogram amounts in normal human and mouse serum. It is increased during parasitic infestations¹ and mediates allergy. CD23, the low-affinity receptor for IgE (FcεRII), has been proposed as an important regulator of IgE synthesis²⁻⁴. The type-II transmembrane lectin⁴ CD23 is expressed in the mouse on B cells and follicular dendritic cells. In humans there are two forms of CD23 which differ in their intracellular amino-terminal 6/7 amino acids⁴; expression of the A-form corresponds to that of murine CD23, whereas the B-form is also found on T and other haematopoietic cells⁴. CD23 has been implicated in cellular adhesion⁵, antigen presentation⁶, as a growth and differentiation factor for human B, T and plasma cells, and as a signal transduction molecule⁷ (reviewed in refs 3, 8). Here we disrupt the gene coding for murine CD23 (ref. 9) to clarify the role of CD23 *in vivo* and find that B- and T-cell development is normal in these CD23-deficient mice. Immune responses to the helminth *Nippostrongylus brasiliensis* are unaffected. In contrast, immuni-

zation with thymus-dependent antigens leads to increased and sustained specific IgE antibody titres compared with controls. Formation of germinal centres is normal. These results suggest that murine CD23 acts as a negative feedback component of IgE regulation.

Our targeting strategy was aimed at inactivating the transmembrane domain encoded by exon 3 of *Fcer2a*¹⁰. Exons 3 and 4 were replaced by a neomycin-resistance (*neo*^r) cassette in a C57BL/6-derived embryonic stem (ES) cell line¹¹ (Fig. 1a). Southern blot analysis showed a proper targeting event (Fig. 1b). The absence of messenger RNA for CD23 and membrane-bound protein was demonstrated by northern blot analysis (Fig. 1c) and flow cytometry (Fig. 1d), respectively.

The composition of the lymphoid compartments and the differentiation of lymphocytes were studied by multi-parameter flow cytometry. The number of B and T cells in peripheral lymphoid organs did not differ between mutant and wild-type mice. Maturation of B cells in the bone marrow was not affected using CD24, CD44, CD45, BP-1, IgD and IgM as differentiation markers. B-cell subpopulations, defined as B1 or B2 subsets, were unchanged. The T-cell compartment in the thymus and the periphery was unaffected, as judged by the expression of the surface markers CD3, CD44, CD24, CD25 and CD69 (data not shown). The absence of CD23 apparently does not impair B- and T-cell development. These results do not support proposals that CD23 is a growth factor for human B and T cells^{12,13}.

In CD23-deficient mice, serum levels of IgM, IgG3 and IgA were not affected. But IgG1 and IgG2b levels were twofold lower, whereas IgE levels were increased twofold as compared with controls. To measure the half-life of IgE, mice were injected with serum obtained from IgE-transgenic mice¹⁴. The rate of disappearance of IgE was identical in CD23-deficient and control mice (Fig. 2). These results indicate that CD23 is not essential for class switch or for plasma-cell differentiation¹⁵ and clearance of serum IgE.

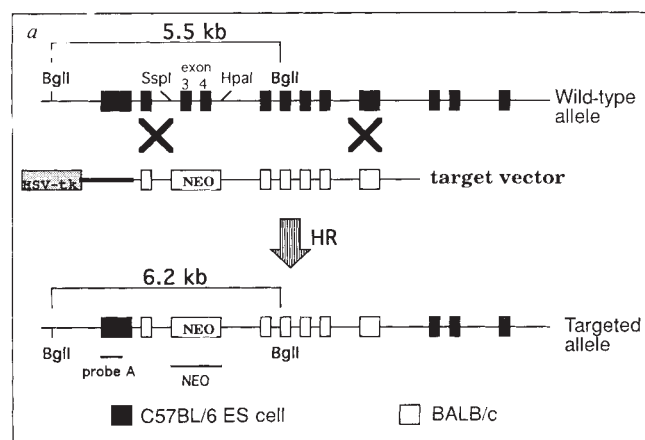
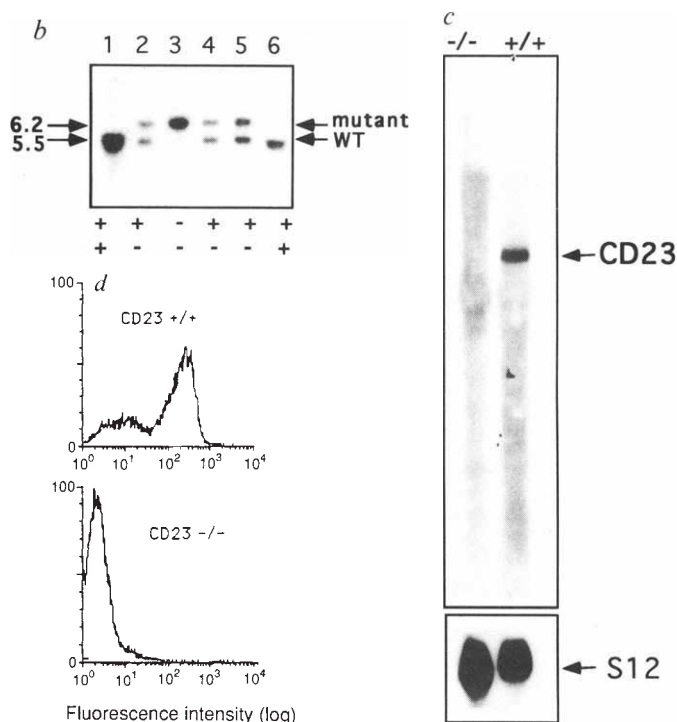
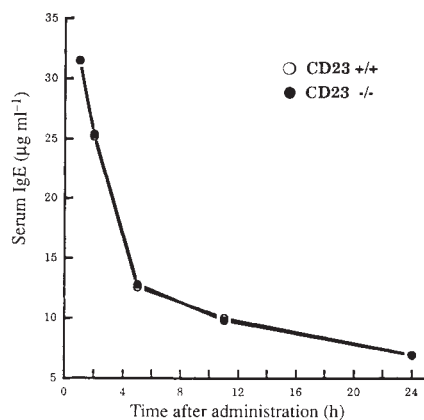


FIG. 1 Gene targeting of *Fcer2a*, the gene coding for murine CD23. a, Strategy for homologous recombination. A 408-base-pair (bp) *SspI*-*HpaI* fragment containing exons 3 and 4 was replaced by a *neo*^r cassette. The homologies on the short arm and the long arm are 1.1 kilobases (kb) and 3.4 kb, respectively. We positioned a herpes simplex virus thymidine kinase (HSV-tk) cassette at the 5' end of the vector. Isolation of clones and screening by PCR has been described²⁷. HR, homologous recombination. b, Southern blot analysis. Genomic DNA from neomycin-resistant ES cell lines or ES-cell-derived offspring was digested with *BglII*. Southern hybridization with probe A confirmed the disruption of the *Fcer2a* gene. The wild-type band is 5.5 kb and the targeted allele 6.2 kb. Lanes 1-4: tail DNA from different progeny (+/+, wild-type (WT); +/-, heterozygous mutant; and -/-, homozygous mutant); lane 5: *Fcer2a*-targeted ES-cell clone ISA 22; lane 6: C57BL/6 ES cell line (not manipulated). c, Northern blot of RNA of homozygous mutant (-/-) and wild-type (+/+) mice. 15 µg total splenic RNA was loaded per lane. Top, the blot was hybridized with a 600-bp fragment that covered coding



sequences from the exons 8 to 12; arrow, 2.2 kb. Bottom, same blot reprobed for the ribosomal protein S12 (ref. 28) to verify the amount of RNA loaded. d, Flow-cytometric analysis of spleen cells. B cells were labelled with antibodies B3B4 (ref. 29; anti-CD23, biotinylated and visualized with streptavidin-phycoerythrin) and 6B2 (anti-CD45 (B220)), a pan B-cell marker, labelled with fluorescein isothiocyanate (FITC). Histograms of B220⁺ cells are shown.

FIG. 2 Clearance of IgE in CD23-deficient and wild-type mice. Mice (2 per group) were injected intravenously with 400 μ l serum of IgE-transgenic mice¹⁴ containing 200 μ g ml⁻¹ IgE. At the indicated times after injection, total serum IgE was measured by enzyme-linked immunosorbent assay (ELISA)¹⁹.



We compared the immune response of CD23-deficient and wild-type mice using three protocols. First, mice were challenged with the thymus-independent antigen 2,4,6-trinitrophenyl-Ficoll (TNP-Ficoll). CD23^{-/-} mice mounted a twofold higher IgM response (not shown). Next, antibody titres were measured following immunization with the thymus-dependent antigen 2,4-dinitrophenyl-ovalbumin (DNP-OVA). Anti-DNP titres for IgM, IgG2b and IgG3 were comparable. In contrast, specific IgG1 was twice as high, and specific IgE 6–12 times as high, in CD23^{-/-} mice (Fig. 3a, b). Total IgE was also raised (Fig. 3c), and in addition, the decrease of IgE in CD23^{-/-} mice was slower (Fig. 3b, c). Similar results were obtained in CD23-deficient mice after immunization with another haptenated antigen, 4-hydroxy-3-nitrophenyl-OVA (NP-OVA; results not shown), and the non-haptenated antigen keyhole limpet haemocyanin (KLH) (Fig. 3d, e).

The formation of germinal centres¹⁵ was investigated in mice immunized with DNP-OVA. Development of germinal centres in CD23^{-/-} mice was not affected (Fig. 4). Interestingly, although

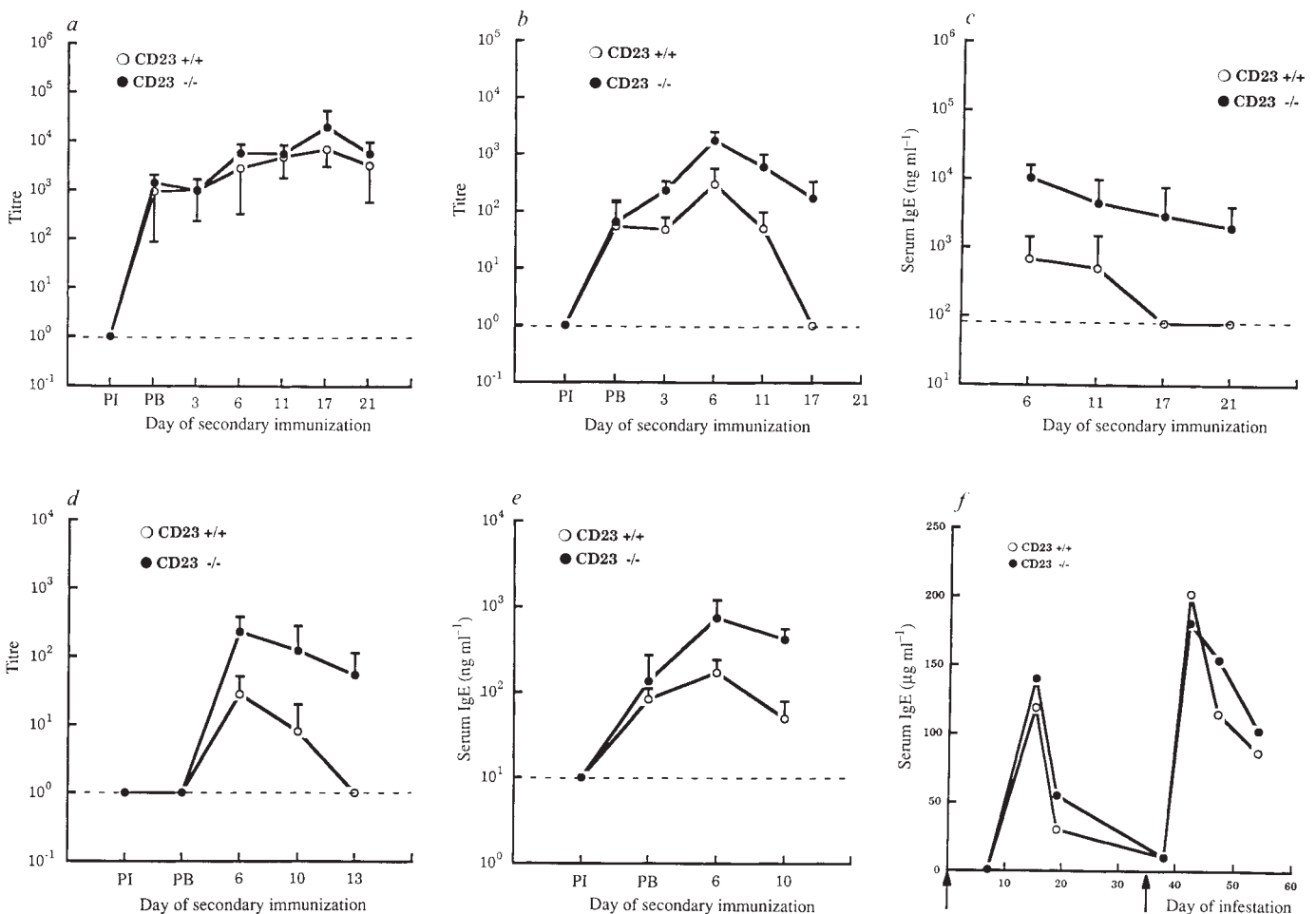


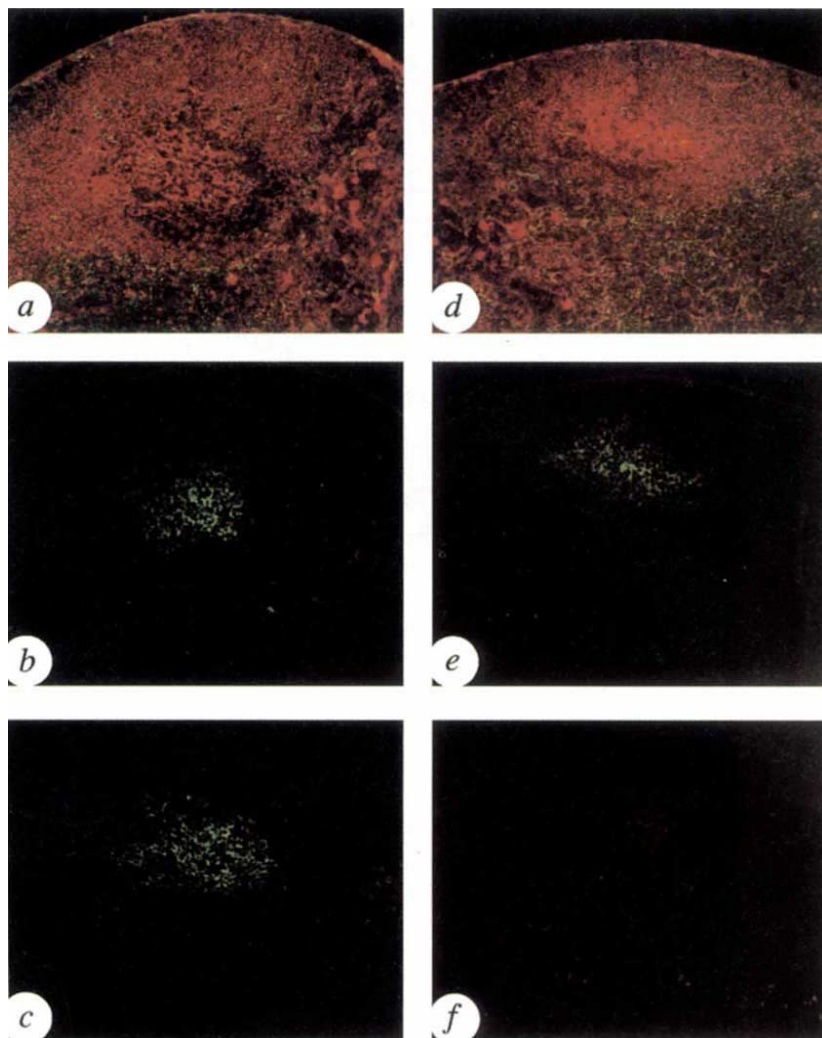
FIG. 3 Immunization experiments and serum antibody levels. a, DNP-specific IgG1; b, DNP-specific IgE; and c, total IgE, in DNP-OVA-immunized mice; d, KLH-specific IgE; e, total IgE in KLH-immunized mice; and f, total IgE in *N. brasiliensis*(Nb)-infested mice. a–e, Mice were immunized subcutaneously (s.c.) with 20 μ g DNP-OVA or 20 μ g KLH (Calbiochem) precipitated in alum with *Bacillus pertussis*. Three weeks later, they were immunized intraperitoneally (i.p.). Serum was taken at different times before and after booster immunizations. Results are expressed as the serum dilution where half-maximum absorbance values were obtained (a, b, d). Serum IgE levels were calculated from serially diluted serum samples (c, e, f). Values from individual mice were used to determine arithmetic means ($\pm$ s.d.). PI, pre-immune; PB, pre-booster. Dashed line represents the lower detection limit. The detection of KLH-specific IgE antibodies was facilitated by depletion of

IgG antibodies: serum was diluted to 200 μ l with PBS buffer and added to 50 μ l packed protein G-Sepharose beads (Pharmacia). After overnight depletion at 4 °C, the supernatant was used for ELISA. Murine IgE did not bind to protein G (manuscript in preparation). f, Mice were immunized s.c. with 700 *N. brasiliensis* larvae stage 3 (Nb L3) at day 0 and again after 35 days (arrows). Serum was taken before and at different times after primary and secondary infestations. Results are given as arithmetic means of serum IgE levels of individual mice.

METHODS. For immunization experiments, 10 eight-week-old animals (DNP-OVA and Nb) or 6 six-week-old animals (KLH) were used per group. Serum levels were determined by ELISA¹⁹. Specific antibodies were determined by ELISA using DNP-bovine serum albumin or KLH as antigen¹⁹.

FIG. 4 Immunohistological evaluation of germinal-centre development, CD23 expression and IgE localization. Serial cryostat sections were prepared from lymph nodes of DNP-OVA-immune mice nine days post-secondary immunization. *a, b* And *c* are from wild-type mice, and *d, e* and *f* are from CD23-deficient mice. *a* And *d* show labelling of IgM-positive B cells in red and Thy1-positive T cells in green. Note the strong red immune-complex labelling of the FDC network positioned adjacent to the IgM-negative area of the germinal centre dark zone. *b* and *e* demonstrate the ability of the FDC to trap immune complexes containing IgE; *c* shows the level of CD23 expression in wild-type mice, and *f* confirms the absence of this molecule in the deficient mouse.

METHODS. 10- μ m-thick frozen sections were prepared from lymph nodes taken 9 days post-secondary challenge with DNP-OVA in alum with *B. pertussis*. The expression of cell-surface molecules was visualized using sheep anti-mouse IgM-Texas Red (Southern Biotechnology), and T24 (anti-Thy1), B3B4 (ref. 29; anti-CD23) or 1-5 (anti-IgE; gift from G. Le Gros) with a mouse anti-rat IgG (specific for heavy and light chain) F(ab')₂ conjugated to FITC reagent (Jackson Immuno Research).



follicular dendritic cells (FDCs) of deficient mice did not express CD23 (Fig. 4*f*), IgE was detected on the surface of these cells (Fig. 4*e*). This indicates that IgE can be bound to FDCs either directly or complexed with other molecules such as complement components and their receptors¹⁶. Our results show that CD23 is not essential in the initiation and maintenance of an immune response. The proposed role of CD23 in antigen presentation^{6,17} is normally of minor importance *in vivo*.

The final immunization protocol involved infestation with the helminth *N. brasiliensis*, which induces strong IgE production¹⁸. Antibody production was similar in both types of mice with respect to total IgM, IgG1, IgG2b, IgG3 and IgA (not shown). CD23^{-/-} mice did not produce significantly more serum IgE than wild-type mice (Fig. 3*f*). Therefore, it seems that CD23 is not involved in primary or secondary parasite-induced antibody responses.

We propose that the function of CD23 lies in a negative feedback control of IgE. It is noteworthy that the interleukin IL-4 is not only essential for class switch to IgE¹⁹ but is also a strong inducer of CD23 expression²⁰. In an IgE-transgenic mouse model it was shown that binding of monomeric IgE to CD23 and subsequent increased expression of CD23 are not sufficient to interfere with antigen-specific IgE responses¹⁴. We therefore postulate that binding of IgE-containing immune complexes to CD23 renders the B cell unresponsive to further stimulation or directly suppresses IgE synthesis. Further support for this model comes from the observations: (1) anti-CD23 antibodies or IgE-containing immune complexes suppress human B-cell proliferation²¹ and IgE production *in vitro*^{22,23}; and (2) anti-CD23 antibody

treatment of rats inhibits a specific IgE immune response². The latter observation was interpreted to favour a role for CD23 as an inducer of IgE synthesis. Alternatively, CD23 might convey a negative signal, in support of our hypothesis. The observation that CD23 does not regulate IgE synthesis in infestation with *N. brasiliensis* can be attributed to the dominant activation of Th-2 cells²⁴ and the strong T-cell-independent, nonspecific activation of B cells¹⁸ by the helminth that overrides the regulatory function of CD23.

In conclusion, murine CD23 is a negative controlling element in IgE production. Its function is reminiscent of that of Fc γ R, which also turns off B cells after binding specific immune complexes^{25,26}. A dysfunction of CD23 may play a part in the aetiology of IgE-mediated diseases such as atopy, allergy and asthma. □

Received 10 November 1993; accepted 22 April 1994.

- Hagan, P., Blumenthal, U. J., Dunn, D., Simpson, A. J. G. & Wilkins, H. A. *Nature* **349**, 243–245 (1991).
- Flores-Romo, L. et al. *Science* **261**, 1038–1041 (1993).
- Conrad, D. H. A. *Rev. Immun.* **8**, 623–645 (1990).
- Sutton, B. J. & Gould, H. J. *Nature* **366**, 421–428 (1993).
- Aubry, J.-P., Pochon, S., Graber, P., Jansen, K. U. & Bonnefoy, J.-Y. *Nature* **358**, 505–507 (1992).
- Kehry, M. R. & Yamashita, L. C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7556–7560 (1989).
- Kolb, J.-P. et al. *J. Immun.* **150**, 4798–4809 (1993).
- Gordon, J. (ed.) *Monogr. Allergy: CD23* **29** (Karger, Basel, 1991).
- Thomas, K. R. & Capocchi, M. R. *Cell* **51**, 503–512 (1987).
- Richards, M. L., Katz, D. H. & Liu, F. T. *J. Immun.* **147**, 1067–1074 (1991).
- Ledermann, B. & Bürki, K. *Expt Cell Res.* **197**, 254–258 (1991).
- Swendeman, S. & Thorley-Lawson, D. A. *EMBO J.* **6**, 1637–1642 (1987).
- Bertho, J.-M., Fourcade, C., Dalloul, A. H., Debré, P. & Mossalayi, M. D. *Eur. J. Immun.* **21**, 1073–1076 (1991).

14. Adamczewski, M., Köhler, G. & Lamers, M. C. *Eur. J. Immun.* **21**, 617–626 (1991).
15. Liu, Y.-J., Johnson, G. D., Gordon, J. & MacLennan, I. C. M. *Immun. Today* **13**, 17–21 (1992).
16. Maeda, K. et al. *J. Immun.* **148**, 2340–2347 (1992).
17. Heyman, B., Tianmin, L. & Gustavsson, S. *Eur. J. Immun.* **23**, 1739–1742 (1993).
18. Lebrun, P. & Spiegelberg, H. L. *J. Immun.* **139**, 1459–1465 (1987).
19. Kopf, M. et al. *Nature* **362**, 245–247 (1993).
20. Keegan, A. D., Snapper, C. M., Van Dusen, R., Paul, W. E. & Conrad, D. H. *J. Immun.* **142**, 3868–3874 (1989).
21. Luo, H., Hofstetter, H., Banchereau, J. & Delespesse, G. *J. Immun.* **146**, 2122–2129 (1991).
22. Sherr, E., Macy, E., Kimata, H., Gilly, M. & Saxon, A. *J. Immun.* **142**, 481–489 (1989).
23. Saxon, A., Kurbe-Leamer, M., Behle, K., Max, E. E. & Zhang, K. *J. Immun.* **147**, 4000–4006 (1991).
24. Street, N. E. et al. *J. Immun.* **144**, 1629–1639 (1990).
25. Campbell, K. A., Lees, A., Finkelman, F. D. & Conrad, D. H. *Eur. J. Immun.* **22**, 2107–2112 (1992).

26. Sinclair, N. R. *StC. Immun. Today* **4**, 35–36 (1983).
27. Nitschke, L., Kopf, M. & Lamers, M. C. *BioTechniques* **14**, 914–916 (1993).
28. Ayane, M. P., Neilsen, P. & Köhler, G. *Nucleic Acids Res.* **17**, 6722 (1990).
29. Rao, M., Lee, W. T. & Conrad, D. H. *J. Immun.* **138**, 1845–1851 (1987).

ACKNOWLEDGEMENTS. We thank H. Mittelstädt, E. Strätling, L. Eggersdorff, C. Westphal and A. Fehrenbach for technical assistance; B. Ledermann for the B6 ES cell line; L. Nitschke and M. Kopf for help with gene targeting; D. Conrad for *Fcer2a* cDNA; and F. McConnell, B. Imhof, P. Shaw, R. Carsetti, M. Simon and D. H. Katz for critical reading of the manuscript. M.R. is a recipient of a Special Fellow Award from the Leukemia Society of America. This work was supported in part by a grant from the NIH to D. H. Katz. The Basel Institute for Immunology was founded and is supported by Hoffman-La Roche, Basel, Switzerland.

A mammalian protein targeted by G1-arresting rapamycin–receptor complex

Eric J. Brown, Mark W. Albers, Tae Bum Shin, Kazuo Ichikawa, Curtis T. Keith, William S. Lane* & Stuart L. Schreiber†

Department of Chemistry, Howard Hughes Medical Institute and
* Harvard Microchemistry Facility, Harvard University,
12 Oxford Street, Cambridge, Massachusetts 02138, USA

† To whom correspondence should be addressed

THE structurally related natural products rapamycin and FK506 bind to the same intracellular receptor, FKBP12, yet the resulting

complexes interfere with distinct signalling pathways^{1,2}. FKBP12–rapamycin inhibits progression through the G1 phase of the cell cycle in osteosarcoma³, liver^{4,5} and T cells^{6,7} as well as in yeast⁸, and interferes with mitogenic signalling pathways that are involved in G1 progression^{9,10}, namely with activation of the protein p70^{S6k} (refs 5, 11–13) and cyclin-dependent kinases^{3,14–16}. Here we isolate a mammalian FKBP–rapamycin-associated protein (FRAP) whose binding to structural variants of rapamycin complexed to FKBP12 correlates with the ability of these ligands to inhibit cell-cycle progression. Peptide sequences from purified bovine FRAP were used to isolate a human cDNA clone that is highly related to the *DRR1/TOR1* and *DRR2/TOR2* gene products from *Saccharomyces cerevisiae*^{8,17,18}. Although it has not been previously demonstrated that either of the *DRR/TOR* gene products can bind the FKBP–rapamycin complex directly^{17,19}, these yeast genes have been genetically linked to a rapamycin-sensitive pathway and are thought to encode lipid kinases^{17–20}.

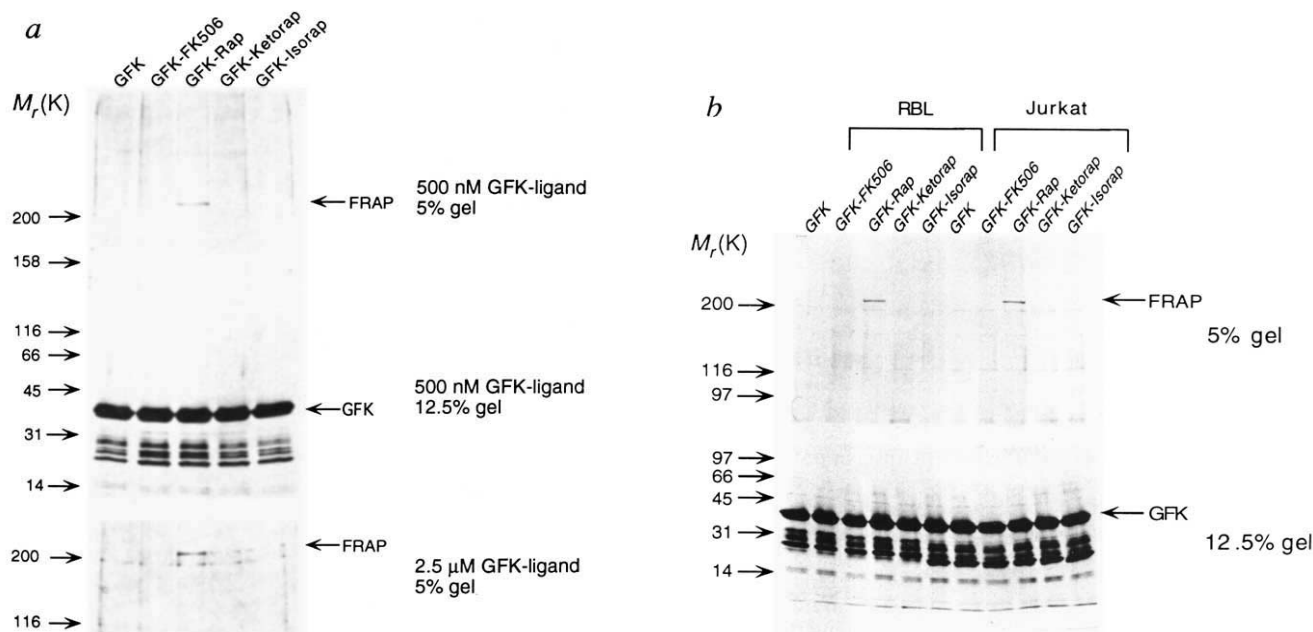


FIG. 1 Identification of FRAP protein in three mammalian cell lines. *a*, GFK alone or individual GFK–ligand complexes were added to MG-63 cell lysates (2×10^7 cells per condition) to a final concentration of either 500 nM or 2.5 μ M and the mixtures incubated for 10 min at 4 °C. Fusion protein complexes were recovered by glutathione-affinity chromatography, and the proteins detected by silver staining after 5% SDS–PAGE. Because of compression, FRAP is not resolved by 12.5% SDS–PAGE, so both 5% and 12.5% gels are shown. The amount of FRAP that was retained by affinity chromatography saturated at concentrations of GFK–Rap greater than 500 nM in these experiments and in others using concentrations of GFK–Rap ranging from 100 nM to 5 μ M (data not shown). *b*, GFK alone or individual GFK–ligand complexes were added to a final concentration of 500 nM to lysates prepared from either 2×10^8 Jurkat T lymphocytes or 10^8 rat basophilic leukaemia (RBL) cells per condition. Lysates were treated as in *a*. FKBP12, but not FKBP13 or FKBP25 (ref. 23) is able to mediate the actions of rapamycin in *S.*

cerevisiae. In addition, we found that YFK188 (ref. 24), an FKBP12 null strain, could be complemented with GFK (P. K. Martin, B. Gladstone, G. Weiss, D. T. Hung, S.L.S., in preparation). Thus the GST appendage of the fusion protein does not preclude binding of the biologically relevant target to the GFK–rapamycin complex in yeast.

METHODS. MG-63, Jurkat and RBL cells were grown in media containing 10% FBS and lysed at 4 °C in PINT buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 2 mM EDTA, 2 mM EGTA, 25 mM NaF, 100 μ M Na_3VO_4 , 25 mM 2-glycerophosphate, 0.2 mM PMSF, 1 μ g ml^{-1} leupeptin, 1 μ g ml^{-1} pepstatin A and 2 mM DTT) containing 0.5% Triton X-100. Lysates were clarified by centrifugation at 25,000g, and the Triton X-100 in the supernatant was diluted to 0.33% by adding 0.5 vol PINT buffer. GFK prebound to stoichiometric quantities of FK506, keto- iso- or unmodified rapamycin was added to lysates as described. Each condition was then passed through a 250- μ l glutathione-Sepharose column, which was washed with PINT buffer containing 0.5 M NaCl and 0.3% Triton X-100.

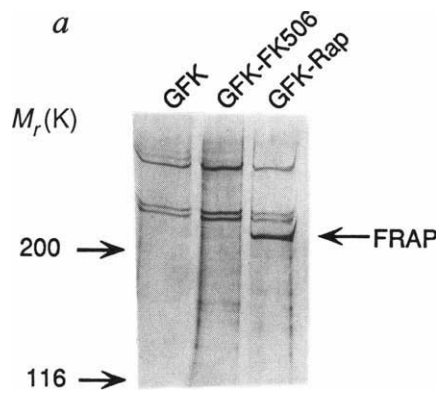


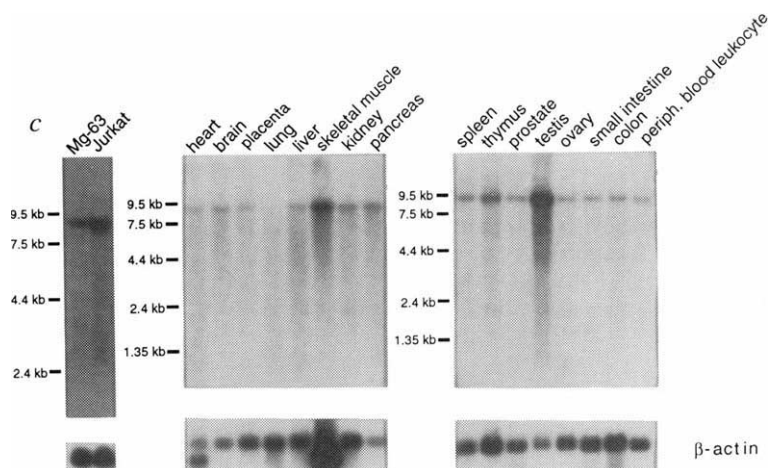
FIG. 2 Purification of FRAP from bovine brain and cDNA cloning of human FRAP. **a**, Fivefold-enriched bovine FRAP (S-column eluate; see below) was conditioned with [100 nM], glutathione-S-transferase-FKBP12 fusion protein (GFK), GFK-FK506 or GFK-Rap. Complexes with fusion proteins were recovered by glutathione-affinity chromatography and detected as described in Fig. 1 legend. We also found FRAP in bovine liver and thymus. **b**, Predicted translational product of the human FRAP cDNA clone. Bovine FRAP peptide sequences aligned to human FRAP are indicated by underlined segments. In the reading frame shown translational stop codons were not encountered upstream of the initiating methionine. **c**, Northern blot analysis of human tissue, Jurkat T cell and MG-63 cell poly (A)⁺ RNA. The Jurkat/MG-63 and multiple tissue Northern blots (Clontech) were hybridized with ³²P-labelled probes derived from the 182 bp PCR fragment and the 5.5 kb clone (text), respectively. Hybridization to human β -actin probe is shown as an internal control for loading.

METHODS. Bovine FRAP was purified by grinding 900 g of bovine brain in blender with 1 litre of PIP (0.3% Triton X-100, 50 mM sodium phosphate, pH 7.2, 2 mM EDTA, 2 mM EGTA, 25 mM NaF, 100 μ M Na₃VO₄, 25 mM 2-glycerophosphate, 1 mM PMSF, 1 μ g ml⁻¹ leupeptin, 1 μ g ml⁻¹ pepstatin A, 1 mM benzimidazole and 2 mM DTT). The homogenate was centrifuged at 25,000g and the supernatant (20 g total protein) was loaded onto a 1 litre S-Sepharose (Pharmacia) column. The column was then washed with PIP and eluted with PINP (PIP with 1M NaCl). GFK-rapamycin was added to the pooled eluate to a final concentration of 100 nM and recovered by glutathione-affinity chromatography. FRAP was resolved by SDS-PAGE and transferred to PVDF. Following digestion with trypsin or endoproteinase Lys-C (Boehringer Mannheim) bFRAP peptides were micro-sequenced²³. The Jurkat T cell cDNA library (Stratagene) was constructed through random and oligo dT priming of cytoplasmic oligo dT purified RNA (ref. 25). cDNA screening, Jurkat and MG-63 RNA isolation and northern blotting and were performed by procedures similar to those previously described²⁵. A 182 bp fragment was amplified from a human brain stem library (Stratagene) and labelled by incorporation of ³²P-dCTP in the course of reamplification by PCR. The sequences were analysed using BLAST (ref. 26) and the University of Wisconsin GCG (ref. 27) software. The human FRAP cDNA sequence has been submitted to Genbank.

We used two structural variants of rapamycin, 16-keto-rapamycin (S. D. Meyer and S.L.S., manuscript in preparation) and 25,26-iso-rapamycin²¹, to identify any biologically relevant targets of the FKBP-rapamycin complex. Both variants bind tightly to human FKBP12, as shown by their ability to inhibit rotamase activity of the recombinant protein (*K_i* values were 0.2 nM for rapamycin⁶, 2 nM for keto-rapamycin, and 0.1 nM for iso-rapamycin). But the variants are about two orders of magnitude less potent than rapamycin in preventing the progression through G1 of MG-63 human osteosarcoma cells. The values of IC₅₀ (half-maximal inhibitory concentration) estimated from dose-response curves are 0.1 nM, 7.5 nM and 50 nM for rapamycin, keto- and iso-rapamycin, respectively. Thus the complexes of iso- and keto-rapamycin with FKBP12 should bind to

b

1	MLGTGPAAT	TAATSSNV	VLOQFASGLK	SRNEETRAKA	AKELQHYVTM	50
51	ELREMSQEE	TRFYDQNH	TEFLVSSDA	NERKGGILAI	ASLIGVEGNN	100
101	ATRIGRFANY	LRNLLPSNDP	VVMEMASKAI	GRLAMAGDTF	TAEYVEFEVK	150
151	RLLEWL GADR	NEGRRAHVAL	VLRELAISVP	TFFQVQVPP	FDNIFVAVWD	200
201	PKQATREGAV	AALRACLILT	TQREPKEMQK	PQWYRHTFEE	AEKGFDETLA	250
251	KEKGMNDRDR	IHGALLILNE	LVRISSEMEG	RLREEMEEIT	QQQLFVHKYC	300
301	KDLMEGFTKP	RHITPTSTFO	AVOPQSSNAL	VGLLGYSSHQ	GLMGFTGSPS	350
351	PAKSTLVESR	CCRDLMEEKF	DQVCQWLK	RNSKNSLIQM	TILNLLPRLA	400
401	AFRPSAFTDT	QYLQDTHNVH	LSVCVKEKER	TAAFQALGLL	SVAVRSEFKV	450
451	YLPRLVDIIR	AALPPKDFAH	KRQKAMQVDA	TVFTCISMLA	RAMGPGIQDD	500
501	IKELLEPMIA	VGLSPALTAV	LYDLSRQTPQ	LKKDQDGLL	KMLSLVLMHK	550
551	PLRHPGMPKG	LAHQLASPLG	TTLPEASDVG	SITLALRTL	SFEFFGSLT	600
601	QFVRHCADHF	LNSEHKIERM	EAARTCSRL	TPSIHLISGH	AHVVSQTAVQ	650
651	VVADVLSKLL	VVGITDPPDP	IRYCVLASLD	RFDAHLAQ	ENLQALFVAL	700
701	NDQVFEIREL	AICTVGRLLS	MNPAFVMPFL	RKMLTQLTE	LEHSGIGRIK	750
751	EQSARMLGHL	VSNAPRLIRP	YMEPILKALI	LKLKDPDPP	NPQVNNVLA	800
801	TIGELAQVSG	LEMKRWDEL	FIIMDMQLD	SSLLAKRQVA	LWTLQQLVAS	850
851	TGVVVEPYRK	YPTLLEVLN	FLKTEQNGQT	RREAIRVLG	LGLADPYKHK	900
901	VNTGMDQSR	DASAVLSSES	KSSQSSSDYS	TSEMLVMNG	LPLDEFFPAV	950
951	SMVALMRIFR	DQSLSHHHTM	VVQAITFIFK	SLGLKCVQFL	PQVMPFTLVN	1000
1001	IRVCDGAIRE	FLFQQLGMLV	SFVKSHIRPY	MDEIVTLMR	FWVMNTSIQS	1050
1051	TIILLIEQIV	VALGGEFKLY	LPLQIPMLR	VFMHNSPGR	IVSILKLAAT	1100
1101	QLFGANLDDY	LHLLLPPIVK	LFDAPLPLP	SRKAALETVD	RLTESLDETD	1150
1151	YASRIIPIIV	RILDQSPELB	STAMDTLSSL	VFLQKKYQI	FIPMNVKVLV	1200
1201	RHRINHORYD	VLCIRIVKGY	TLADEEEDFL	TYOHRMLRSQ	OGDALASGPV	1250
1251	ETGPMKKLHV	STINLOKAWG	AABRVSKDDW	LEWRLRLSLE	LLKDS SSPSL	1300
1301	RSCWALAQAY	NPMARDLFNA	AFVSCWSELN	EDQDELIRS	TELALTSQDI	1350
1351	AEVTQTLNL	AEFMEHSDKG	PLPLRDNDGI	VLLGERAAKC	RAYAKALHYK	1400
1401	ELEEQKGTP	AILESILSIN	NKLQPEAA	GVLEYAMKHE	GELEIQATWY	1450
1451	EKLHEWEDAL	VAYDKKMDTN	KDDPELMLGR	MRCLEALGEW	QGLHQCECEK	1500
1501	WTLVNDTEQA	KMARMAAAAA	WGLQWDSME	EYTCMPROT	HDSGAFYRAVL	1550
1551	ALHQDLFSLA	QOCIDKARDL	LDDEL TAMAG	ESYSRAGVAM	ESYCHMLSE	1600
1601	EVQYKLVPE	RREIRIQIWW	ERLQGCORIV	EDWQKILMVR	SLVVS PHEDM	1650
1651	RTHLYASL	GKSGRLALAH	KTLVLLGLVD	PSRQLDPLP	TVHPQVITYAY	1700
1701	MKNMWSARK	IDAFQHMQH	VQTMQQAQH	ATATEDQHQH	QELHKL MARC	1750
1751	FLKLGEQWLN	LQGINESTIP	KVLQYYSAT	EHDRSWYKAW	HAWAYMNFEE	1800
1801	VLHYKHQNOA	RDEKKLRLHA	SGANITNATT	AATTAATATT	TASTEGNSE	1850
1851	SEASTENSP	TPSPLOKQVT	EDLSKTLMY	TPPAVQGFRR	SISLSRGNNL	1900
1901	QDTLRVLTW	FDYGHWPDVN	EALVEGVKAI	QDITWLVQIP	QLIARIDTPR	1950
1951	PLVGRILHQL	LTDIRYHPQ	ALIYPTVAS	KSTTTARHNA	ANKILKNMCE	2000
2001	HNSTLVQQA	MYSEELIRVA	ILWHEMMHEG	LEASRLYFG	ERNVKGMEV	2050
2051	LEPLHAMMER	GPQTLKEISE	NOAYGDLME	AQEWCKRYMK	SGNVKDLTQA	2100
2101	WDLVYHFR	ISKQLPQLTS	LELOVYSPKL	LRCRLLELV	PGTYDNPQPI	2150
2151	IRTSIAPSL	QVITSKQRP	KLTLMSGNG	EFVFLKXGHE	DLRQDERVMQ	2200
2201	LFGLVNTLLA	NDPSTLRKNI	STORYAVIPL	STNSGLIGWV	PHCDTLHALI	2250
2251	RDYREKKKIL	LNIEHRIMLR	MAPQYDHLT	MQKVVEFEHA	VNNTAGDRLA	2300
2301	KLLWLKSPSS	EWVDRRTNY	TRSLAVMSMV	GYLGLGDRH	PSNLMRLDLS	2350
2351	GKTLHIDFGD	CFEAVMTREK	FPEKIPFRLT	RMLTNAMETV	GLDGNRYTIC	2400
2401	HTVMEVLEH	KDSVMALVEA	FVYDPLNWR	LMDNTKGNK	RSRTRTDSYS	2450
2451	AGOSVFELDG	VELGEPAHKK	TGTTVPESIH	SEFGDGLVFC	FALNKKATQI	2500
2501	INRVRDKLTG	RDFSDDTLD	VPTQVELLIK	QATSHENLCQ	FYIGWCPFW	2549



the FKBP12-rapamycin target less effectively than FKBP12-rapamycin itself.

A fusion protein of glutathione-S-transferase with FKBP12 (GFK) was used to identify candidates for the biologically relevant targets of FKBP12-rapamycin. MG-63 cells were lysed by detergent and complexes of GFK-rapamycin, GFK-FK506 or GFK alone were added individually to clarified lysate at a final concentration of 500 nM or 2.5 μ M (Fig. 1a). A protein of approximate relative molecular mass 220,000 (*M_r* ~ 220K) was detected in the GFK-rapamycin sample by SDS-PAGE and silver staining (Fig. 1a, lane 3). This FKBP-rapamycin-associated protein (FRAP) was not retained with GFK-FK506 or GFK alone (Fig. 1a, lanes 1 and 2). No other rapamycin-specific proteins were detected by silver staining (Fig. 1a) or by a similar

affinity purification procedure using lysates from [³⁵S] methionine-labelled cells (data not shown). The GFK-ketorapamycin and GFK-isorapamycin complexes bound FRAP less effectively than GFK-rapamycin; at concentrations of 500 nM, the keto- and iso- complexes were unable to retain the 220K protein (Fig. 1a, lanes 4, 5), whereas at higher concentrations of the complexes (2.5 µM) detectable quantities of FRAP were retained (Fig. 1a, lanes 4, 5). This is consistent with the finding that these compounds are still strong cell-cycle inhibitors, albeit less potent than rapamycin itself. Thus, the binding of GFK-ligand complexes to FRAP correlates with the ability of the ligands to impede G1 progression in MG-63 cells. FRAP was also detected in Jurkat T-lymphocyte cells and rat basophilic leukaemia cells (Fig. 1b), two mammalian cell lines that are also sensitive to rapamycin^{6,22}. No other rapamycin-specific bands were observed in each case.

FRAP purified from bovine brain (bFRAP) had a similar specificity for GFK-ligand (Fig. 2a). Microsequencing of bFRAP proteolytic fragments (298 amino acids in total, Fig. 2b) led to the design of a pair of degenerate oligonucleotides for use in the polymerase chain reaction (PCR). A 182 bp PCR product allowed for the isolation of overlapping clones from a human Jurkat T cell λZAP II cDNA library, yielding 7.6 kb of contiguous sequence. Using these cDNA sequences as probes, a band migrating at approximately 8.5 kilobases was detected by Northern blot analysis of oligo dT purified RNA isolated from a variety of human tissues and cell lines (Fig. 2c). The human cDNA sequence encodes an amino-acid open reading frame (ORF) and aligns with 99% identity to the bFRAP peptides (Fig. 2b). As N-terminal peptide sequence from purified bovine FRAP was not obtained, the initiating methionine shown in Fig. 2b is unconfirmed. The predicted molecular mass of this ORF (~300K) is greater than that inferred by the mobility of FRAP during SDS-PAGE (above).

Human FRAP is highly related to the *DDR1/TOR1* and *DDR2/TOR2* gene products. Overall it is 44% identical to *DDR1/TOR1* and 46% identical to *DDR2/TOR2*. The region of greatest homology to *DDR1/TOR1* and *DDR2/TOR2* lies in the C-terminal 660 amino acids of human FRAP (57% and 59% identical, respectively). In addition, this region has homology to several known phosphatidylinositol kinases (21% identity on average), including mammalian phosphatidylinositol 3-kinase^{17,18} (PI3K), a yeast PI3K *VPS34* (refs. 17 and 18) and *PIK1* (ref. 20). These similarities indicate that FRAP may also have phosphatidylinositol kinase activity.

Through the introduction of minute structural changes in rapamycin, this study implicates FRAP as a mediator of G1 cell cycle progression in mammalian cells. Identification of FRAP as the target of FKBP12-rapamycin together with the earlier demonstration of calcineurin as the target of FKBP12-FK506 (ref. 2) addresses a fascinating aspect of immunophilin research, namely that the immunophilin FKBP12 can bind two distinct natural products and thereby gain the ability to bind two distinct signalling molecules involved in cell cycle entry and progression. Further biochemical characterization of this unique mammalian protein should elucidate its role in propagating the mitogen-initiated signals that lead to the activation of p70^{S₆} and cyclin-Cdk complexes. □

Received 3 May; accepted 20 May 1994.

- Schreiber, S. L. *Science* **251**, 283–287 (1991).
- Schreiber, S. L. *Cell* **70**, 365–368 (1992).
- Albers, M. W. et al. *Ann. N. Y. Acad. Sci.* **696**, 54–62 (1993).
- Francavilla, A. et al. *Hepatology* **15**, 871–877 (1992).
- Price, D. J., Grove, J. R., Calvo, V., Avruch, J. & Bierer, B. E. *Science* **257**, 973–977 (1992).
- Bierer, B. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9231–9235 (1990).
- Dumont, F. J., Staruch, M. J., Koprak, S. L., Molino, M. R. & Sigal, N. H. *J. Immun.* **144**, 251–258 (1990).
- Heitman, J., Movva, N. R. & Hall, M. N. *Science* **253**, 905–909 (1991).
- Lane, H. A., Fernandez, A., Lamb, N. J. C. & Thomas, G. *Nature* **363**, 170–172 (1993).
- Norbury C. & Nurse, P. A. *Rev. Biochem.* **61**, 441–470 (1992).
- Chung, J., Kuo, C. J., Crabtree, G. R. & Blenis, J. *Cell* **69**, 1227–1236 (1992).
- Kuo, C. J. et al. *Nature* **358**, 70–73 (1992).

- Calvo, V., Crews, C. M., Vik, T. A. & Bierer, B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7571–7575 (1992).
- Morice, W. G., Wiederrecht, G., Brunn, G. J., Siekierka, J. J. & Abraham, R. T. *J. biol. Chem.* **268**, 22737–22745 (1993).
- Morice, W. G., Brunn, G. J., Wiederrecht, G., Siekierka, J. J. & Abraham, R. T. *J. biol. Chem.* **268**, 3734–3738 (1993).
- Albers, M. W. et al. *J. biol. Chem.* **268**, 22825–22829 (1993).
- Cafferkey, R. et al. *Molec. cell. Biol.* **13**, 6012–6023 (1993).
- Kunz, J. et al. *Cell* **73**, 585–596 (1993).
- Helliwell, S. B. et al. *Molec. Biol. Cell* **5**, 105–118 (1994).
- Flanagan, C. A. et al. *Science* **262**, 1444–1448 (1993).
- Hayward, C. M., Yohannes, D. & Danishefsky, S. J. *J. Am. chem. Soc.* **115**, 9345–9346 (1993).
- Hultsch, T., Martin, R. & Hohman, R. J. *Molec. Biol. Cell* **3**, 981–987 (1992).
- Galat, A., Lane, W. S., Standaert, R. F. & Schreiber, S. L. *Biochemistry* **31**, 2427–2434 (1992).
- Foor, F. et al. *Nature* **360**, 682–684 (1992).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning A Laboratory Manual*, 2nd edn (Cold Spring Harbor Laboratory Press, New York, 1989).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. molec. Biol.* **215**, 403–410 (1990).
- Devereux, J., Jaegerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387–395 (1984).

ACKNOWLEDGEMENTS. M.W.A. and E.J.B. have made an equal contribution to this research. We thank S. J. Danishefsky and C. M. Hayward for isorapamycin; S. D. Meyer for preparing ketorap; R. F. Standaert for advice and assistance; P. A. Beal for critically reading the manuscript; Anjana Rao and Lewis Kunkel for helpful discussions; and R. Robinson, R. Chic and V. Bailey for their assistance. This work was supported by a grant from the NIH (S.L.S.). M.W.A. is a Howard Hughes Medical Institute predoctoral fellow, E.J.B. is an NIH predoctoral fellow, K.I. is a visiting scientist sponsored by Sankyo Co. Ltd. and S.L.S. is an Investigator with the Howard Hughes Medical Institute.

Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex

Dimitris Tzamarias & Kevin Struhl*

Department of Biological Chemistry and Molecular Pharmacology, 240 Longwood Avenue, Harvard Medical School, Boston, Massachusetts 02115, USA

DNA-BINDING repressor proteins mediate regulation of yeast genes by cell type (*Mcm1/a2* and *a1/a2*), glucose (*Mig1*) and oxygen (*Rox1*) (refs 1–4 respectively). An unusual feature of all these regulatory pathways is that transcriptional repression requires two physically associated proteins⁵ that do not bind DNA *Cyc8*(*Ssn6*) and *Tup1*. The *Cyc8-Tup1* complex has been proposed to be a co-repressor that is recruited to target promoters by pathway-specific DNA-binding proteins⁶, but the specific functions of the individual proteins are unknown. Here we show that when it is bound upstream of a functional promoter through the LexA DNA-binding domain, *Tup1* represses transcription in the absence of *Cyc8*. Deletion analysis indicates that *Tup1* contains at least two non-overlapping transcriptional repression regions with minimal primary sequence similarity, and a separable *Cyc8*-interaction domain. These *Tup1* domains, which do not include the β-transducin motifs⁷, are necessary and partially sufficient for *Tup1* function. We suggest that *Tup1* performs the repression function of the *Cyc8-Tup1* co-repressor complex, and that *Cyc8* serves as a link with the pathway-specific DNA-binding proteins.

It has been previously shown that *Cyc8* can repress transcription in a *Tup1*-dependent manner when bound upstream of the intact *CYC1* promoter through the heterologous LexA DNA-binding domain⁶. Similarly, a LexA-*Tup1* hybrid protein confers a 16-fold reduction of expression from a promoter containing four LexA operators upstream of the *CYC1* promoter (Table 1). LexA-*Tup1* and LexA-*Cyc8* also repress expression of a *his3* gene containing a single LexA operator upstream of the T_R TATA element (Fig. 1a), suggesting that they can inhibit basal transcription. Surprisingly, LexA-*Tup1* retains almost its entire

* To whom correspondence should be addressed.

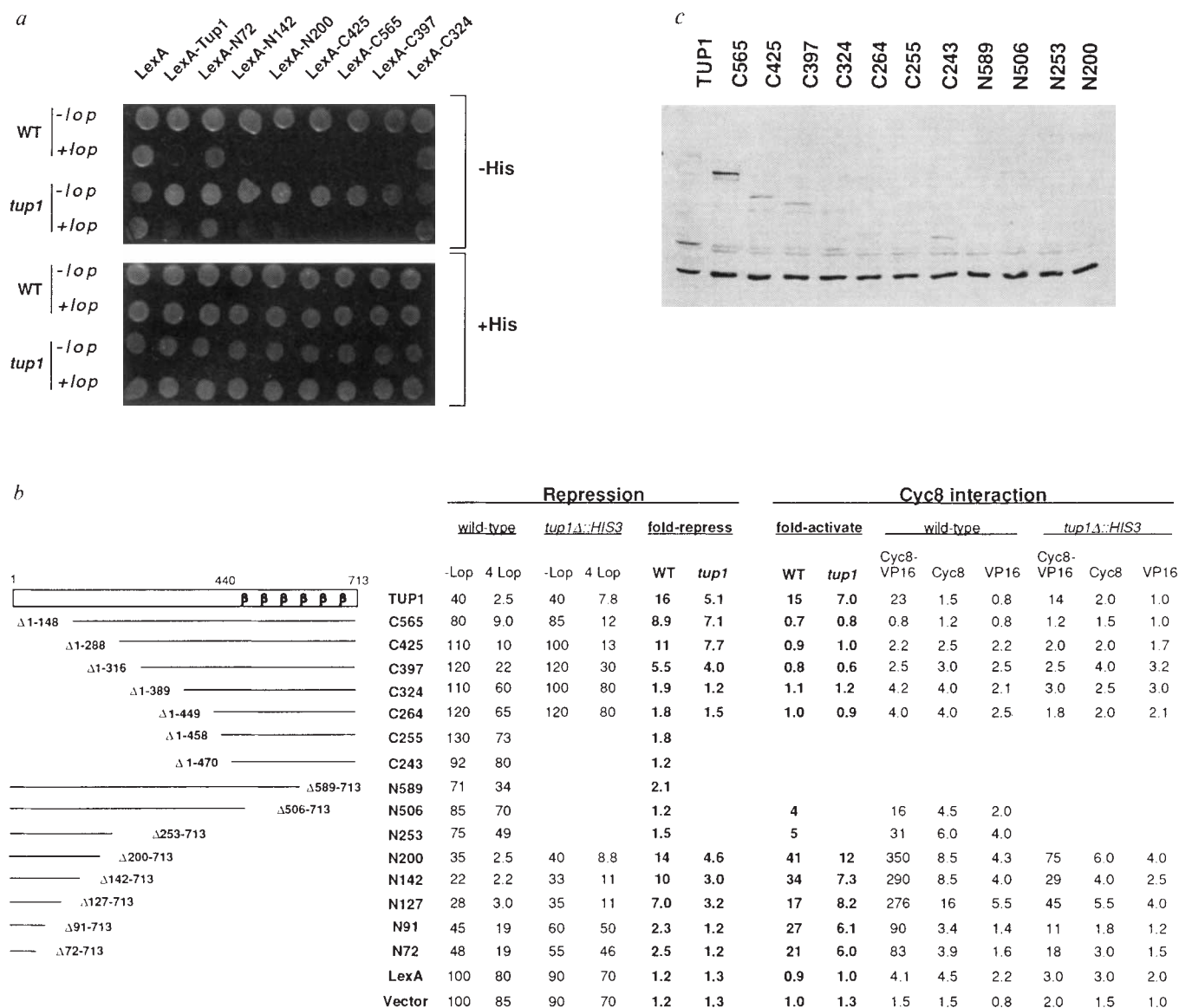
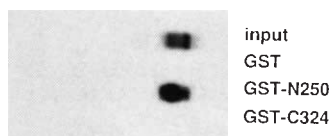


FIG. 1 Mapping the transcriptional repression and Cyc8-interaction domains of Tup1. **a**, Repression of basal transcription. YCp91 plasmids expressing various LexA-Tup1 derivatives were introduced into isogenic wild-type and *tup1Δ::HIS3* strains containing a *his3* promoter with (+lop) or without (–lop) a single LexA operator upstream of the *T_R* TATA element; the resulting cells were plated on minimal medium in the presence (+) or absence (–) of histidine (His). Repression of *HIS3* transcription was assayed by the inability of the resulting cells to grow on –His plates. LexA-Cyc8 also represses basal transcription in this assay (data not shown). **b**, Transcriptional repression and 2-hybrid assays. The structures of Tup1 (713 amino acids, including six copies of a β -transducin motif) and deletion derivatives (named by the number of N- or C-terminal residues contained in the protein) are indicated together with the sequences that have been removed. Repression assays were carried out by measuring β -galactosidase activities of wild-type and *tup1Δ::HIS3* strains harbouring the indicated LexA-Tup1 derivatives and *LacZ* reporter constructs that do or do not contain 4 LexA operators upstream of the *CYC1* promoter (Table 1 legends). The decreased *LacZ* expression from the *CYC1* promoter due to the various LexA-Tup1 proteins correlates well with the degree of Tup1 function and slower cell growth of the strains. 2-hybrid assays were carried out by measuring β -galactosidase activities of wild-type and *tup1Δ::HIS3* strains expressing the indicated combinations of LexA-Tup1 derivatives and activation (Cyc8-VP16) or control (Cyc8 or VP16) proteins. The strains all contain a plasmid derived from JK103 (ref. 22) in which *LacZ* is controlled by a promoter containing 4 LexA operators upstream of the *GAL1* TATA element. The ability of each LexA-Tup1 derivative to interact with Cyc8 is determined by the fold-activation (ratio of activity with Cyc8-VP16 compared to Cyc8 control). Weak activation of the N253 derivative might reflect lowered protein levels rather than functional activity. We do not understand why repression and two-hybrid interactions of all derivatives tested are reduced (typically ~3-fold) in the *tup1Δ::HIS3* strain in comparison to the isogenic wild-type strain. **c**, Western analysis of LexA-Tup1 derivatives. Electrophoretically separated proteins from yeast strains containing the indicated proteins were probed with the HA1 monoclonal antibody²³, which recognizes the 'flu epitope located at the N terminus of each protein. Bands

appearing in all lanes represent yeast proteins that react with the HA1 antibody; they are observed in the parental strain lacking any LexA-Tup1 protein.

METHODS. *CYC8* and *TUP1* sequences were obtained from yeast genomic DNA by PCR amplification using primers based on published sequences^{7,24}. The YCp91 expression vector is a derivative of the *TRP1* centromeric plasmid pRS314 (ref. 25) that contains the *ADH1* promoter and 5' untranslated region (nucleotides –410 to +10), followed by an ATG codon, sequences encoding the SV40 nuclear localization signal²⁶, the HA1 epitope from influenza virus²³, a polylinker, and the *CYC8* termination region (includes 410 bp beyond the stop codon). To express the LexA hybrid proteins, an *EcoRI* fragment containing the entire LexA coding region (residues 1–202) was inserted upstream of the nuclear localization sequence. LexA-Cyc8 contains the entire Cyc8 coding region including 9 bp from the 5' untranslated region cloned between the *Bam*HI and *Nco*I sites, and LexA-Tup1 contains the Tup1 coding region cloned between the *Sma*I and *Kpn*I sites. The LexA-Tup1 deletion derivatives were made by *Bal*31 nuclease treatment of the Tup1 coding region with subsequent insertion into the *Sma*I and *Kpn*I sites of YCp91. The activation and control constructs used in the 2-hybrid experiment were cloned into YEp92, which was generated by inserting the entire expression cassette of the YCp91 (*Sph*I–*Sac*I fragment) into the 2 μ /LEU2 plasmid YEpLac181 (ref. 27). The Cyc8-VP16 activation construct contains the *Bam*HI–*Asp*718 fragment of *CYC8* (codons 1–300) fused to the *Bgl*III–*Bam*HI fragment harbouring the VP16 activation domain (codons 414–553)²⁸, whereas the control constructs contain the individual *CYC8* or VP16 fragments. The portion of Cyc8 in the hybrid protein complements a *cyc8* deletion allele²⁰ and represses transcription in a Tup1-dependent manner when bound to a promoter through the LexA DNA-binding domain (data not shown). The *his3* promoters used in the basal transcription experiment was generated by inserting GGATCCGCATACCAACCATTAACCCCTACTGATGTACATACAGTAGTGTGG-TCACAGAAATGGATCC (Lex operator underlined) into the *Bam*HI site of a derivative containing a mutated GCN4 site upstream of the *T_R* element²⁹; the distance between the LexA operator and the *T_R* TATA element is 45 bp.

FIG. 2 Cyc8-Tup1 interaction *in vitro*. Each lane represents 35 S-labelled Cyc8 (residues 1–351, which contains 9 TRP motifs and complements a cyc8 deletion allele²⁰) stably bound to agarose beads containing GST, GST-Tup1-N250, and GST-Tup1-C324. The lane representing input 35 S-Cyc8 contains only 20% of the amount that was incubated with the agarose beads.



METHODS. The *Bam*HI–*Mlu*I fragment of Cyc8 was cloned downstream of the T3 promoter, and 35 S-protein was synthesized by transcription and translation *in vitro* using T3 RNA polymerase and rabbit reticulocyte lysates. The plasmids expressing GST-N250 and GST-C324 were generated respectively by cloning the *Sma*I–*Bam*HI fragment of Tup1 or the *Sma*I–*Nco*I fragment from C324 into the pGEX-2T vector, and the proteins were bound to glutathione agarose as described³⁰. Agarose beads (50 μ l) containing 1 μ g of the GST proteins were pre-incubated for 30 min at 4 °C with 0.25% BSA and then incubated with 35 S-Cyc8 for 2 h in binding buffer (20 mM Tris, pH 7.5, 100 mM NaCl, 0.1% NP-40, 2 mM PMSF, 0.25% BSA). Beads were washed 3 times with binding buffer (25 column volumes each) by microcentrifugation, and once with binding buffer lacking BSA. Samples were analysed by SDS-PAGE.

repression activity in a *cyc8* deletion. Thus unlike repression by LexA–Cyc8, which requires Tup1 (ref. 6), repression by LexA–Tup1 does not require Cyc8. This suggests that Tup1 mediates the repression function of the Cyc8–Tup1 co-repressor complex.

Deletion analysis of LexA–Tup1 indicates that at least two non-overlapping regions in the N-terminal half of Tup1 can mediate the repression function (Fig. 1b). The derivative containing the N-terminal 200 residues (N200) represses transcription to the same extent as LexA–Tup1, and the N142 and N127 derivatives are nearly as efficient; N91 and N72 are not functional. However, derivatives lacking the N-terminal 148 (C565) or 288 (C425) residues also repress transcription, although less efficiently than LexA–Tup1. C397 partially represses transcription, whereas more truncated derivatives do not, even though the hybrid proteins are expressed at levels comparable to LexA–Tup1 (Fig. 1c). In all cases, repression occurs in a *tup1* deletion strain, but for unknown reasons, the effect is less pronounced. Thus, the 200 N-terminal residues and a region(s) C-terminal to residue 288 can independently mediate repression.

To identify the region of Tup1 that associates with Cyc8, we used the 'two-hybrid' method for defining protein protein interactions⁸. We introduced the various LexA–Tup1 derivatives into strains expressing a Cyc8 VP16 hybrid protein. As Cyc8 and Tup1 are physically associated⁵, the combination of Cyc8 VP16 and LexA–Tup1 confers 15-fold more expression than either protein alone (Fig. 1b). Similar results are obtained with LexA hybrid proteins containing 72, 91, 127 and 200 N-terminal residues of Tup1. These two-hybrid interactions are also observed to a lesser extent in a strain containing the *tup1* Δ :*HIS3* allele. Conversely, derivatives lacking 148 or more

N-terminal residues do not stimulate transcription above the LexA control molecule. These results suggest that the N-terminal 72 residues of Tup1 are necessary and sufficient for the formation of the Cyc8–Tup1 complex. The Cyc8 interaction and transcriptional repression functions are separable because the N91 and N72 derivatives do not mediate repression even though they interact with Cyc8.

Affinity chromatography using glutathione-S-transferase (GST) fusion proteins suggests that the N-terminal region of Tup1 interacts directly with Cyc8. A radiolabelled Cyc8 derivative binds tightly to a GST column containing Tup1-N250, but it fails to bind to GST–Tup1-C324 (which contains the β -transducin repeats) or to a GST column (Fig. 2).

We also examined the Tup1 deletion derivatives (LexA sequence removed) for two natural functions of the Cyc8–Tup1 complex, repression of *SUC2*, a glucose-regulated gene, and repression of *ANB1*, an oxygen-regulated gene (Fig. 3). N200, which contains the Cyc8 interaction and repression functions, reduces both *SUC2* and *ANB1* transcription, and it largely rescues the temperature-sensitive growth and clumpy-colony phenotypes caused by a *tup1* deletion. Repression of *SUC2* and *ANB1* by N200 requires Cyc8 (data not shown) but is not complete. In contrast, all derivatives lacking the Cyc8 interaction region (C565 to C243) fail to repress *SUC2* and *ANB1* transcription, even though some of them can mediate LexA-dependent repression. The N72 derivative, which interacts with Cyc8 but does not mediate LexA-dependent repression, also fails to reduce *SUC2* and *ANB1* transcription. These observations indicate that the Cyc8 interaction and transcriptional repression domain are necessary and partially sufficient for Tup1 function.

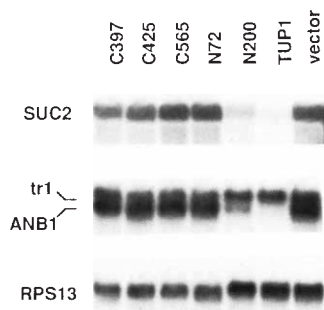
Our results suggest that Tup1 contains a domain that is responsible for the transcriptional repression function. Within this domain, short non-overlapping regions with minimal sequence similarity can independently mediate the repression

TABLE 1 Transcriptional repression by LexA–Tup1 and LexA–Cyc8 hybrid proteins

Protein	Strain	Promoter		Fold-repression
		–Lex op.	4 Lex op.	
LexA–Tup1	Wild type	40	2.5	16.0
LexA–Tup1	<i>cyc8</i> Δ 9::HIS3	34	2.7	12.6
LexA–Tup1	<i>tup1</i> Δ 1::HIS3	40	7.8	5.1
LexA–Tup1	<i>tup1</i> Δ 1::HIS3 <i>cyc8</i> Δ 9::LEU2	36	6.9	5.2
LexA–Cyc8	Wild type	62	2.3	26.9
LexA–Cyc8	<i>tup1</i> Δ 1::HIS3	27	10	2.7
LexA	Wild type	100	80	1.2
None	Wild type	100	85	1.2

β -galactosidase activities (average of 3 independent transformants) of yeast strains containing the indicated LexA proteins (see legend to Fig. 1), *CYC8* and *TUP1* alleles, and *LacZ* reporter constructs. Values are normalized to the absorbance at 600 nm (A_{600}) of cells at the time of collecting (clumped cells were dispersed by addition of EDTA to 2 mM) and are accurate to $\pm 30\%$. The fold-repression is determined by the ratio of β -galactosidase activities driven by *CYC1* promoters (–324 to +141) that lack (–Lex op.) or contain four Lex A operators (4 Lex op.) 50 bp upstream of the *CYC1* UASs. Cells expressing LexA–Tup1 and LexA–Cyc8 grow slowly and contain only 30–50% of the total protein in control cells. Hence, the apparent decrease in expression from the *CYC1* promoter due to LexA–Tup1 and LexA–Cyc8 proteins reflects normalization of β -galactosidase activities to cell absorbance; when normalized to total protein, all six strains show similar levels of expression from the *CYC1* promoter. The promoter constructs were derived from pLGA312S (ref. 19) and pJK1621 (ref. 6) by removing the *Hind*III fragment that contains the 2 μ origin, and they were integrated at the *URA3* locus. All yeast strains were derived from FT5 (*ura3*–52 *trp1*– Δ 63 *his3*– Δ 200 *leu2*::*PET56*). The *cyc8* Δ 9::HIS3 and *cyc8* Δ 9::LEU2 alleles contain the *Pst*I *HIS3* or *LEU2* fragments between two *Pst*I sites in *CYC8* (codons 99–862) and are essentially identical to *ssn6*– Δ 9 (ref. 20). The *tup1* Δ 1::HIS3 allele contains Sc4251, a *Bam*HI–*Eco*RI *HIS3* fragment²¹, between an artificial *Bam*HI site 6 bp upstream of the ATG initiation codon and the *Eco*RI sites of *TUP1* (codons –2 to 672).

FIG. 3 Repression of *SUC2* and *ANB1* by Tup1 proteins. Total RNAs from *tup1* Δ :*HIS3* deletion strains transformed with the indicated derivatives were separated in a 1.4% agarose gel, transferred to nitrocellulose, and hybridized with radio-labelled *SUC2*, *ANB1* and *RPS13* probes prepared by nick-translation. The *ANB1* probe also hybridizes to TR1, a related gene that is not regulated by oxygen or by the Cyc8–Tup1 complex.



function. These properties are similar to those of transcriptional activation domains, and they suggest that the Tup1 repression domain interacts with a component(s) of the transcriptional machinery. The Tup1 repression domain is almost completely uncharged, and hence differs significantly from artificial repression domains selected from *Escherichia coli* sequences, which are highly basic⁹. Interestingly, it contains two alanine-rich regions (40% of residues 120–144 and 22% of residues 254–276) like those found in the repression domains of the *Drosophila* DNA-binding proteins, Kruppel¹⁰, Engrailed^{11,12} and Even-skipped¹³.

We suggest that Cyc8 may serve as an adaptor between DNA-bound proteins (such as Mig1 and Rox1) and Tup1. Cyc8 and the Cyc8-interaction region of Tup1 are dispensable when the Tup1 repression function is artificially tethered to a promoter, but both are required for Tup1 to repress transcription of *SUC2* and *ANB1*, which is probably mediated by binding of Mig1 and Rox1 to the respective promoters. Conversely, Cyc8 does not repress transcription in the absence of Tup1, and the Cyc8-interaction region of Tup1 is insufficient for repression. Consistent with this adaptor hypothesis, we have made Cyc8 deletion derivatives that interact with Tup1 but differentially repress the pathway-specific genes. This suggests that Cyc8 might interact differently with the various DNA-binding proteins that mediate pathway-specific repression. Our results do not exclude the involvement of Tup1 and/or additional factors in the putative⁶ association of the Cyc8–Tup1 complex with DNA-binding proteins.

Tup1 contains six copies of a 'β-transducin' motif⁷ which has been proposed¹⁴ to interact with proteins containing 'TPR' motifs¹⁵ such as Cyc8. But the properties of the N200 protein indicate that the β-transducin repeats are not absolutely required for Tup1 function. Interestingly several C-terminal deletions that remove one or more β-transducin motifs (N589, N506) are functionally impaired even though they contain the repression and Cyc8-interaction domains (Fig. 1; ref. 7). Thus, when truncated, the C-terminal region of Tup1 can interfere with the repression function and, to a lesser extent, the Cyc8 interaction function. This interference could reflect intramolecular masking or formation of non-productive complexes with other proteins, as appears to be the case with steroid receptors containing partial deletions of the hormone-response domain^{16–18}. Although the β-transducin motifs are not essential for repression or interaction with Cyc8, they affect Tup1 function and are likely to be important. □

Received 12 October 1993; accepted 12 May 1994.

1. Keleher, C. A., Goutte, C. & Johnson, A. D. *Cell* **53**, 927–936 (1988).
2. Goutte, C. & Johnson, A. D. *Cell* **52**, 875–882 (1988).
3. Nehlin, J. O., Carlberg, M. & Ronne, H. *EMBO J.* **10**, 3373–3377 (1991).
4. Zitomer, R. S. & Lowry, C. V. *Microbiol. Rev.* **56**, 1–11 (1992).
5. Williams, F. E., Varanasi, U. & Trumbly, R. J. *Molec. cell. Biol.* **11**, 3307–3316 (1991).
6. Keleher, C. A., Redd, M. J., Schultz, J., Carlson, M. & Johnson, A. D. *Cell* **68**, 709–719 (1992).
7. Williams, F. E. & Trumbly, R. J. *Molec. cell. Biol.* **10**, 6500–6511 (1990).
8. Fields, S. & Song, O.-K. *Nature* **340**, 245–246 (1989).
9. Saha, S., Brickman, J. M., Lehming, N. & Ptashne, M. *Nature* **363**, 648–652 (1993).
10. Licht, J. D., Gossell, M. J., Figge, J. & Hansen, U. M. *Nature* **346**, 76–79 (1990).
11. Jaynes, J. B. & O'Farrell, P. H. *EMBO J.* **10**, 1427–1433 (1991).
12. Han, K. & Manley, J. L. *EMBO J.* **12**, 2723–2733 (1993).
13. Han, K. & Manley, J. L. *Genes Dev.* **7**, 491–503 (1993).
14. Goebel, M. & Yanagida, M. *Trends biochem. Sci.* **16**, 173–177 (1991).
15. Sikorski, R. S., Boguski, M. S., Goebel, M. & Hieter, P. *Cell* **60**, 307–317 (1990).
16. Godowski, P. J., Rusconi, S., Miesfeld, R. & Yamamoto, K. R. *Nature* **325**, 365–368 (1987).
17. Hollenberg, S. M., Giguere, V., Seigui, P. & Evans, R. M. *Cell* **49**, 39–46 (1987).
18. Howard, K. J., Holley, S. J., Yamamoto, K. R. & Distelhorst, C. W. *J. biol. Chem.* **265**, 11928–11935 (1990).
19. Guarente, L. & Mason, T. *Cell* **32**, 1279–1286 (1983).
20. Schultz, J., Marshall-Carlson, L. & Carlson, M. *Molec. cell. Biol.* **10**, 4744–4756 (1990).
21. Tzamarias, D., Pu, W. T. & Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2007–2011 (1992).
22. Kamens, J., Richardson, P., Mosialos, G., Brent, R. & Gilmore, T. *Molec. cell. Biol.* **10**, 2840–2847 (1990).
23. Field, J. et al. *Molec. cell. Biol.* **8**, 2159–2165 (1988).
24. Schultz, J. & Carlson, M. *Molec. cell. Biol.* **7**, 3637–3645 (1987).
25. Sikorski, R. S. & Hieter, P. *Genetics* **122**, 19–27 (1989).
26. Nelson, M. & Silver, P. *Molec. cell. Biol.* **9**, 384–389 (1989).
27. Gietz, R. D. & Sugino, A. *Gene* **74**, 527–534 (1988).
28. Cress, W. D. & Triezenberg, S. J. *Science* **251**, 87–90 (1991).

29. Kim, J., Tzamarias, D., Ellenberger, T. E., Harrison, S. C. & Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **90**, 4513–4517 (1993).
30. Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1988).

ACKNOWLEDGEMENTS. We thank A. Bortvin, F. Winston, M. Collart, J. Kamens, A. Zervos and R. Brent for yeast strains and DNAs, and A. Johnson, M. Monastriotti, M. Oettinger and G. Threos for discussion. This work was supported by a postdoctoral fellowship from the Human Frontiers of Science Program (D.T.) and by a research grant from the NIH.

Three-dimensional structure of β-galactosidase from *E. coli*

R. H. Jacobson, X.-J. Zhang, R. F. DuBose* & B. W. Matthews†

Institute of Molecular Biology, Howard Hughes Medical Institute, and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

THE β-galactosidase from *Escherichia coli* was instrumental in the development of the operon model¹, and today is one of the most commonly used enzymes in molecular biology. Here we report the structure of this protein and show that it is a tetramer with 222-point symmetry. The 1,023-amino-acid polypeptide chain^{2,3} folds into five sequential domains, with an extended segment at the amino terminus. The participation of this amino-terminal segment in a subunit interface, coupled with the observation that each active site is made up of elements from two different subunits, provides a structural rationale for the phenomenon of α-complementation. The structure represents the longest polypeptide chain for which an atomic structure has been determined. Our results show that it is possible successfully to study non-viral protein crystals with unit cell dimensions in excess of 500 Å and with relative molecular masses in the region of 2,000K per asymmetric unit. Non-crystallographic symmetry averaging proved to be a very powerful tool in the structure determination, as has been shown in other contexts^{31,32}.

Crystals of *E. coli* β-galactosidase with four tetramers (each of M_r 465,412K) per asymmetric unit (space group $P2_1$; $a = 107.9$ Å, $b = 207.5$ Å, $c = 509.9$ Å, $\beta = 94.7^\circ$) were obtained as described⁴. To improve the multiple isomorphous replacement phases, which were of poor quality (Table 1), iterative cycles of 16-fold averaging and reconstruction were carried out initially at 5 Å resolution and ultimately extending to 3.5 Å resolution. The resulting map (Fig. 1a) exhibited extremely good connectivity throughout the entire polypeptide chain, was easily interpretable, and was consistent with the amino-acid sequence^{2,3}. The current, partially refined⁵, model has good geometry (Table 1), and includes all residues for the sixteen independent chains, as well as 148 water molecules and two bound magnesium ions per monomer. Representative electron density for the active site region of monomer A is shown in Fig. 1b.

The crystal structure shows β-galactosidase to be a 222-point symmetric tetramer with dimensions of roughly $175 \times 135 \times 90$ Å along the respective two-fold axes. The constituent monomers form two different monomer–monomer contacts that we refer to as the 'activating' interface and the 'long' interface. The 'activating' interface, seen in Fig. 2a as the region of contact between the green/blue (or red/yellow) dimers, includes contacts between residues near the amino termini, and also includes two helices from each monomer that are packed together to form a four-helix bundle (Fig. 2b). In addition, as discussed below, an extended loop reaches across this interface to complete the active site of the neighbouring monomer (Fig. 2b). The 'long' interface is formed by contacts between red/green and blue/yellow dimers. This interface consists of two regions of contact, one of which includes residues near the middle of the sequence and

* Present address: Immunex Corporation, 51 University Street, Seattle, WA 98101, USA.

† To whom correspondence should be addressed.

the other involves contacts within the C-terminal fifth of the sequence.

The β -galactosidase monomer (Fig. 2c) is composed of five compact domains plus ~50 residues at the N-terminus that are relatively extended and contribute to the activating interface. This segment corresponds to the α -complementation peptide and will be discussed separately. The first domain (Figs 2c, 3) corresponds to a 'jelly-roll' β -barrel. The second domain has a fibronectin-type III fold⁶, a topology that is similar to immunoglobulin constant domains. The connection between the fourth

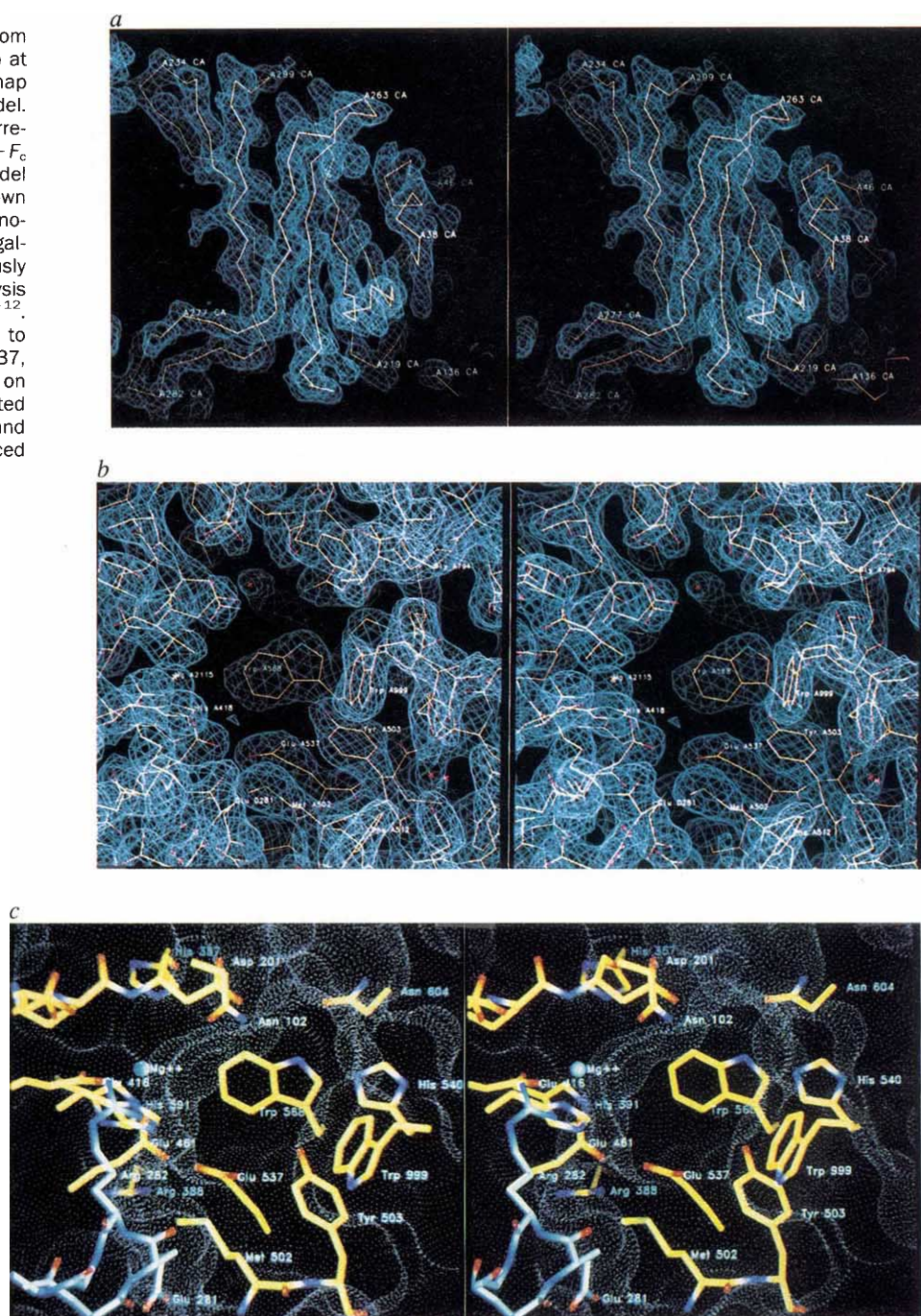
and third strands of the barrel (residues 272–288) forms a protruding loop that extends across the activating interface to the neighbouring monomer (Fig. 2b). The third domain forms a distorted 'TIM' barrel⁷ and contains the catalytic site (see below). Unlike typical TIM barrels that consist of eight β/α repeats, the α/β barrel in β -galactosidase lacks the fifth helix, and the sixth parallel strand of the barrel is distorted. This irregularity results in a space or opening that is occupied by the loop that extends from the second domain of the two-fold related monomer (Fig. 2b) and forms part of the active site (see below).

TABLE 1 X-ray structure determination

Form of crystal	Native (CHESS)	Native (PF)	Hg-c-Gal	TiCl ₃	EMTS
Data statistics					
Crystals	18	8	2	3	4
Soaking time (h)			48	72	72
Soaking conc. (mM)			7.5	Saturated	0.1
Measured reflections	466,195	1,321,660	419,475	386,031	658,596
Unique reflections	199,493	559,917	242,054	238,615	316,782
R_{merge} (%)	9.7	7.0	8.2	8.1	7.7
R_{iso} 3.5 Å (%)			20.4	14.7	12.4
Data to 3.5 Å (%)	70	92	82	73	87
Data to 2.5 Å (%)		73			
Number of sites			36	16	16
R.m.s. F_H /residual (5.0 Å)			1.02	0.80	1.22
R.m.s. F_H /residual (3.5 Å)			0.75	0.58	0.98
R_c (5.0 Å)			0.64	0.67	0.55
R_c (3.5 Å)			0.69	0.70	0.61
Refinement statistics					
Resolution (Å)		∞ –2.5			
Overall R -factor		0.193			
Number of reflections		559,710			
Number of protein atoms		131,168			
Number of solvent atoms		1,981			
R.m.s. bond deviation from ideal (Å)		0.015			
R.m.s. angle deviation from ideal (deg)		2.4			
R.m.s. B (Å ²)		5.0			

Initial X-ray data were collected at the Cornell High Energy Synchrotron Source⁴. All data used in the MIR analysis were from the Photon Factory beamline BL6A2. One native and three derivative data sets were collected using a Weissenberg camera with 860-mm cassette²⁰. Data were recorded on Fuji phosphor-imaging plates and scanned with a Fuji BA100 scanner. Images were transferred to exabyte tape for processing in-house using 'WEIS'²¹ running on a Stardent (Kubota Pacific) Titan. Images were integrated to a resolution where the average $I/\sigma(I)$ value was 2.0, corresponding to d -spacings between 2.0 and 2.9 Å. A similar protocol was followed for the derivative data sets, where the maximum resolution was 3.0–3.5 Å. Heavy-atom derivatives were screened in-house using low-resolution (5 Å) oscillation photography on a Rigaku rotating anode X-ray generator equipped with Charles Supper double-focusing mirrors at a 200-mm crystal-to-film distance. Three potential derivatives were identified: Hg-c-Gal, a mercurated substrate analogue (synthesized by B. Branchaud, L. Blaszcak), TiCl₃, and EMTS (ethyl mercury-thiosalicylate). The Hg-c-Gal and EMTS soaks were in 15% PEG 8000, 60 mM sodium cacodylate, 50 mM NaCl, 80 mM MgSO₄, pH 5.9. The TiCl₃ derivative was prepared by first transferring crystals to mother liquor containing 15% PEG 8000, 60 mM cacodylate, 50 mM NaCl, 100 mM Na₂SO₄, 10 mM EDTA, pH 5.9, and allowing them to equilibrate for 3 days. Crystals were then transferred to 5% PEG 8000, 60 mM cacodylate, 50 mM NaCl, 100 mM Na₂SO₄, saturated TiCl₃, pH 5.9, and allowed to equilibrate for a further 72 h. Twenty initial heavy-atom positions were located in the Hg-c-Gal derivative using vector verification with the program VERIFY (S. L. Roderick, unpublished). These were used to calculate a difference Fourier map that suggested the locations of 16 additional sites. Following refinement of the heavy-atom positions²², phases were calculated at 5 Å and used to solve the EMTS and TiCl₃ derivatives using cross-phased difference Fourier maps. Solutions were verified by analysis of the difference Patterson functions. The thermal parameters of the heavy atoms were all set to 30 Å² and the occupancies and positions were refined²². A 5 Å MIR map (figure of merit 0.54) was calculated using the phases from the three derivatives. Tetramer positions were identified by analysis of the heavy-atom positions and transformations relating the four tetramers to a 160 × 160 × 160 Å orthorhombic pseudo-cell were determined. The electron density within spheres of radius 80 Å centred at each of the 222-point symmetry centres was cut out of the MIR map. Five cycles of averaging and reconstruction of the cell were carried out at 5 Å resolution using the weighting method in ref. 23. Transformations were then refined by maximizing the correlation between the original MIR density and the phase-refined density of the tetramer in the pseudo cell. Twelve cycles of further averaging and reconstruction were carried out at 3.5 Å resolution, resulting in an overall correlation between calculated and observed structure factors of 0.87. The map was readily interpretable except for ~70 residues (729–777, 824–837 and 857–863) that extended outside the 80 Å sphere within which the map had been averaged. Characteristic density corresponding to large aromatic residues throughout the sequence could be readily identified, allowing the unambiguous assignment of the protein sequence. A model for one monomer was built and expanded by the non-crystallographic symmetry operators to generate the full asymmetric unit. An $F_o - F_c$ difference map calculated with phases from this model and averaged clearly showed the locations of all missing residues. Preliminary refinement was carried out using non-crystallographic symmetry constraints in the 'TNT' package of programs (ref. 5, and D. E. Tronrud, unpublished results). An averaged $F_o - F_c$ difference map then identified two magnesium ions and 148 solvent molecules per monomer that were expanded to their equivalent positions in each of the 16 monomers. The coordinates have been deposited in the Brookhaven Data Bank. $R_{\text{merge}} = \sum_i |I_i - \bar{I}| / \sum_i I_i \times 100$, where the sum is taken over the subset of reflections with multiple observations and $\bar{I}$ is the mean value of the i^{th} intensity, I_i . $R_{\text{iso}} = \sum |I_{\text{PH}} - I_P| / \sum I_P \times 100$ where I_{PH} and I_P are the heavy-atom and native intensities. R.m.s. F_H /residual is defined as the mean value of the heavy-atom structure factor divided by the residual lack-of-closure error. $R_c = \sum |F_{\text{PH}} \pm F_P| - F_{\text{H(Calc)}}| / \sum |F_{\text{PH}} - F_P|$, where F_{PH} , F_P and $F_{\text{H(Calc)}}$ are the derivative, native and calculated heavy-atom structure factors, respectively. R -factor = $\sum_{\text{hkl}} |F_{\text{obs}} - F_{\text{calc}}| / \sum_{\text{hkl}} |F_{\text{obs}}|$. R.m.s. B is the r.m.s. discrepancy in the difference in B -values of bonded pairs relative to a standard library of such differences (ref. 5, and D. E. Tronrud, unpublished results).

FIG. 1 a, Stereo view of typical density from the 16-fold averaged, phase-refined, map at 3.5 Å resolution. Superimposed on the map is the α -carbon trace from the initial model. The map is contoured at 2.0σ and corresponds to the second domain. b, $2F_o - F_c$ electron density phased on the refined model and contoured at 1.25σ . Density is shown from the active site region of the first monomer (monomer A). c, The active site of β -galactosidase. Residues previously demonstrated to be important for catalysis are shown (Glu 461, Glu 537, Tyr 503)^{10,12}. Glu 416, His 418 and Glu 461 appear to coordinate a bound magnesium ion. Glu 537, identified as the nucleophile¹¹, is situated on the opposite side of the cavity and is oriented through hydrogen bonds with Tyr 503 and Arg 388. Figure prepared with AVS (Advanced Visual Systems, Inc.).



The fourth domain, residues 628–736, is topologically identical to the second domain. The core of the fifth domain consists of a novel 18-stranded, antiparallel sandwich. From the first to the last strand in each sheet, the direction rotates by over 90°. In addition, the domain also contains an irregular assembly of segments that pack against the back of the first β -sheet and position Trp 999 at the active site.

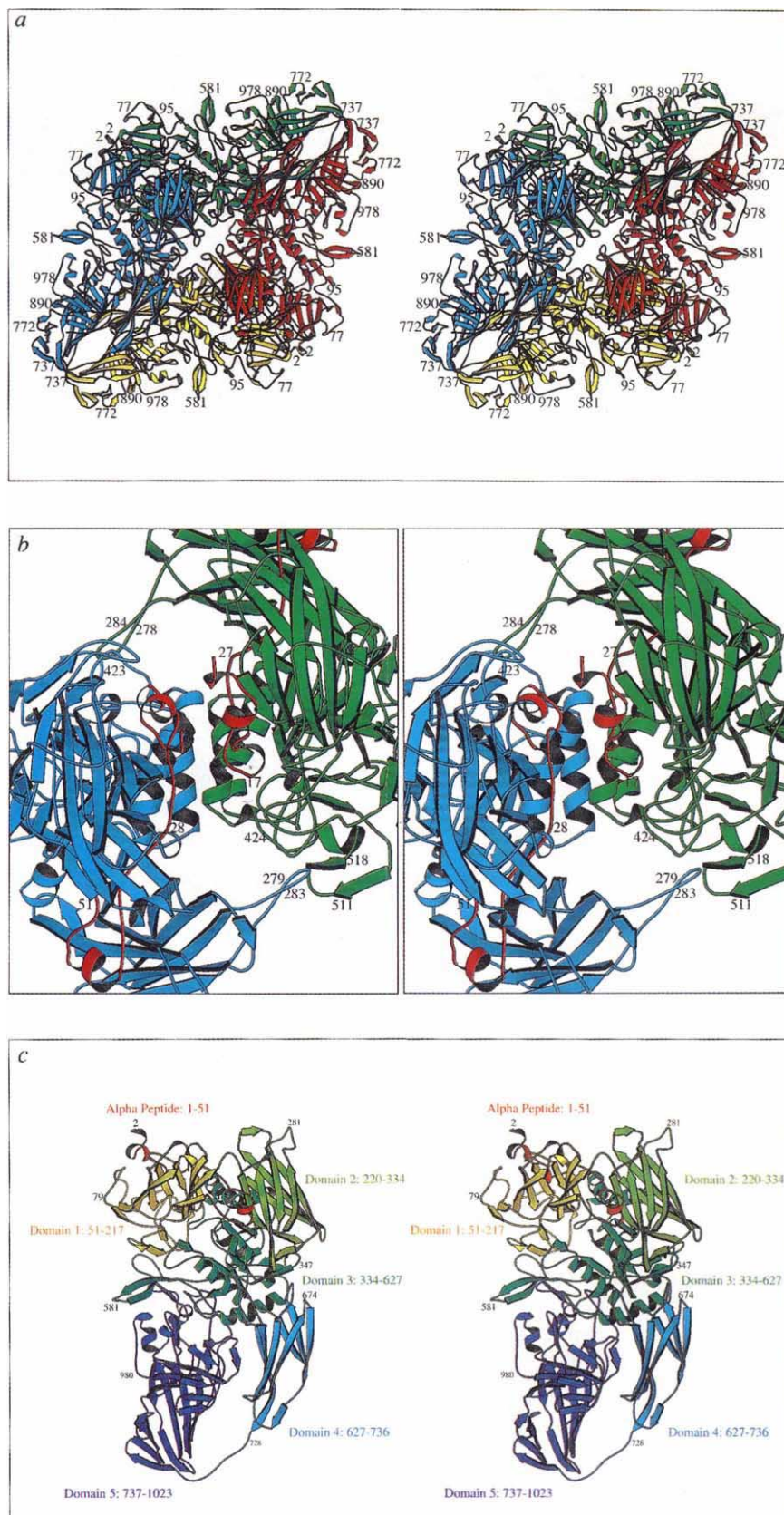
The modular structure of the monomer strongly suggests that the folding of this rather long polypeptide chain is facilitated by each domain acting as an independent folding unit^{8,9}. Indeed, early studies had shown that the C-terminal third of β -galactosidase could fold independently and complement molecules missing this part of the sequence (ω -complementation)⁹. The fifth domain probably corresponds to this so-called ω -fragment.

The residues Glu 461, Met 502, Tyr 503 and Glu 537 have been shown to be important for catalytic function, or near the active site^{10–12}. These residues are found close together and are located around a deep pit formed at the end of the TIM barrel

(Fig. 1b, c). This pocket is therefore identified as the substrate binding site. Its location at the C termini of the β -sheet strands is consistent with other structures that contain the TIM-barrel motif⁷. Among homologous β -galactosidase sequences (Fig. 3), the residues that form the active-site pocket are highly conserved. The loop that extends across the activating interface (Fig. 2b) also contributes to the integrity of the active site. It does not seem to provide catalytic residues *per se*, but to stabilize the main chain in the vicinity of the Mg^{2+} -binding ligands.

α -Complementation¹³ is a property of β -galactosidase that is the basis for the commonly used blue/white screen to identify recombinant DNA and which was first observed in extracts of *E. coli* with deletions in the *lacZ* gene (which encodes β -galactosidase). One of the deletion mutants, the ' α -acceptor' M15, coded for a β -galactosidase lacking residues 11–41 (ref. 14). The resultant protein was inactive and was also dimeric. The second component contained an N-terminal segment, the ' α -donor' or 'complementation peptide'. The α -donor was also inactive, but

FIG. 2 *a*, Ribbon representation²⁴ of the β -galactosidase tetramer showing the largest face of the molecule. Contacts between red/green and blue/yellow dimers form the long interface. Contacts between the red/yellow and blue/green dimers form the activating interface. Formation of the tetrameric particle results in two deep clefts that run across opposite faces of the molecule and each contain two active sites. *b*, Ribbon diagram of the blue/green dimer viewed down the molecular two-fold axis, showing the composition of the activating interface. Residues 1–50 from each chain, which form the α -complementation region (see text), are shown in red. The interface includes contacts between the respective complementation peptides, between two helices from the respective monomers that pack together to form a four-helix bundle, and between an extended loop (residues 272–288) from each monomer that reaches across the interface and extends into the active-site region of the neighbouring monomer stabilizing the active site structure. *c*, Stereo ribbon diagram of the β -galactosidase monomer showing the domain organization of the chain. Residues corresponding to successive domains are coloured in successive spectral colours. The TIM barrel (domain 3) forms the core of the monomer and is surrounded by the four other largely β structures. The extended loop (residues 272–288) that reaches across the activating interface is seen at the upper right.



when mixed with the α -acceptor it 'complemented' to form a protein that was tetrameric and active. Subsequently, many α -donors of varying lengths have been shown to be functional for complementation. The minimum requirement appears to be a peptide containing residues 3–41 (refs 14–17).

The crystal structure of β -galactosidase allows complementation to be rationalized in terms of the formation of the activating

interface which, in turn, allows the protruding loop from domain 2 (residues 272–288) to extend across the activating interface (Fig. 2*b*) and to complete the formation of the active site in domain 3 of the 2-fold-related monomer. The first 50 residues (the 'complementation region') participate in contacts with domains 1, 2 and 3 of the same monomer, as well as with the 2-fold-related complementation region of the monomer across

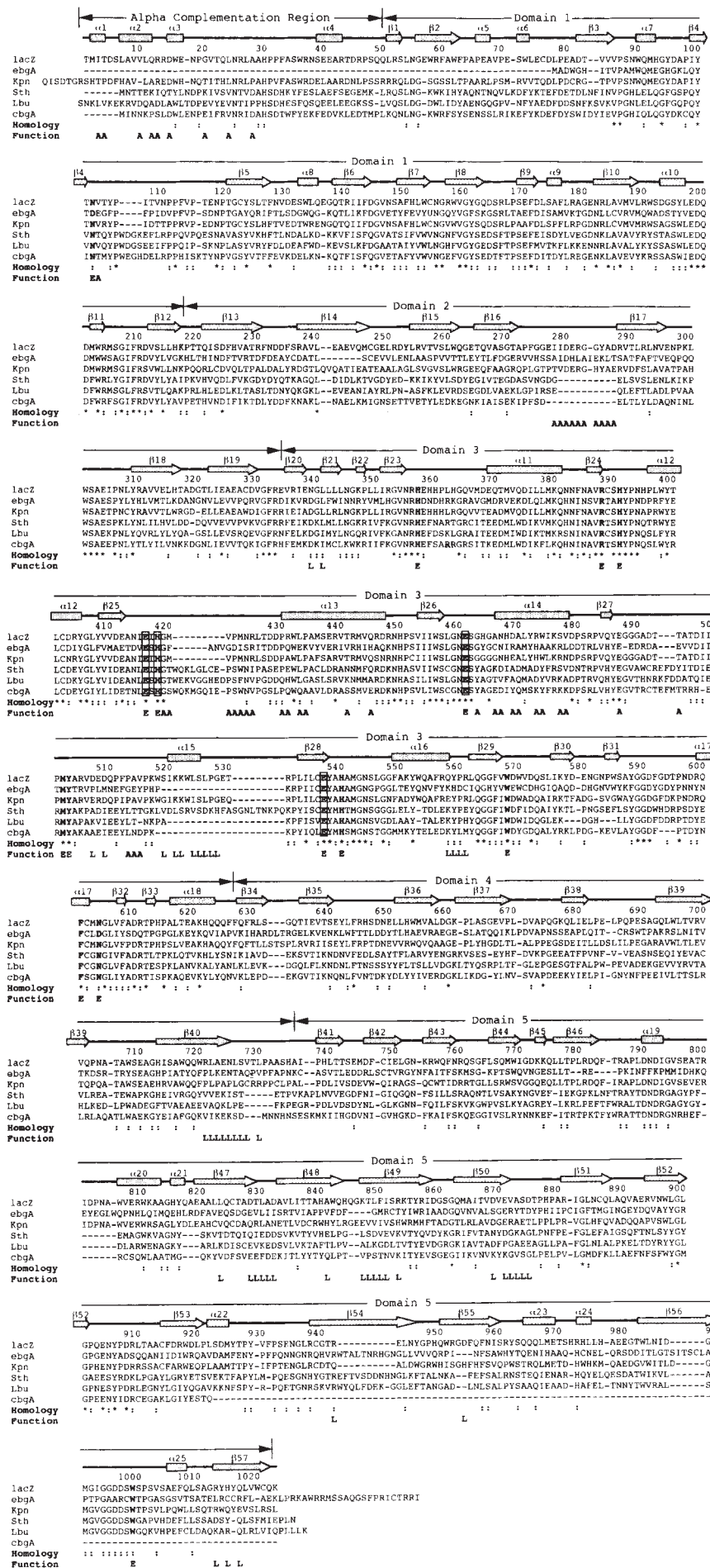


FIG. 3 Domain structure, secondary structure as assigned by DSSP²⁵, and alignment of the sequence of the *E. coli* enzyme with five homologous β -galactosidases. Aligned sequences include: *lacZ*, *E. coli*^{2,3}; *ebgA*, evolved β -galactosidase *E. coli*²⁶; *kpn*, *Klebsiella pneumoniae*²⁷; *sth*, *Streptococcus thermophilus*²⁸; *lbu*, *Lactobacillus bulgaricus*²⁹; *cbgA*, *Clostridium acetobutylicum*³⁰. β -strands are denoted by arrows, α -helices by rectangles. Symbols: (*), absolutely conserved residue; (:), 4 or 5 identical occurrences; (A), residue forming the activating interface (one or more atoms within the residue are closer than 4.5 Å to the 2-fold-related monomer at the activating interface); (L), residues forming the long interface (one or more atoms within the residue are closer than 4.5 Å to the 2-fold related monomer at the long interface); (E), residues forming the substrate-binding pocket and magnesium-binding site (these residues are also shown in bold). The putative nucleophile (Glu 537) and active-site magnesium ligands are boxed.

the activating interface^{15,16}. This interface is formed largely by a four-helix bundle that includes helices $\alpha 13$ and $\alpha 14$ of the TIM barrel (together with their 2-fold-related symmetry mates) (Fig. 2b). The complementation peptide contributes to the interface in two different ways: first, it participates directly in the interfacial contact region; second, it helps indirectly by stabilizing the four-helix bundle.

Different β -galactosidases (Fig. 3) have poor sequence conservation in the vicinity of both the long and the activating interfaces. In particular, the Sth, Lbu and CbgA sequences have a 10–12-residue deletion within the donated loop, and insertions in the vicinities of residues 420 and 493. These changes near the active-site region and the four-helix bundle suggest that the Sth, Lbu and CbgA may lack an activating interface or that the active site might be formed within a single monomeric chain. Consistent with this, Lbu β -galactosidase appears to be dimeric¹⁸.

Why is *E. coli* β -galactosidase so large? Lysozymes cleave the same glycosidic bond but are $\sim 1/20$ th the size. The natural substrates for lysozyme, bacterial cell walls, are highly crosslinked peptidoglycans, so perhaps lysozyme needs to be small to access susceptible bonds¹⁹; β -galactosidase is subject to no such restriction. Inspection of the structure suggests that the different parts combine together in an interdependent way to form the active enzyme. Different regions contribute to the active site, to the long interface, or to the activating interface (Fig. 3). There is no large segment that could be removed without having a deleterious effect. It is conceivable that the protein arose from different structural domains whose combined activity happened to be useful to the cell, and became trapped in sequence space at roughly its present size. □

Received 5 April; accepted 20 May 1994.

- Jacob, F. & Monod, J. *J. molec. Biol.* **3**, 318–356 (1961).
- Fowler, A. & Zabin, I. *J. biol. Chem.* **253**, 5521–5525 (1978).
- Kainins, A., Otto, K., Ruther, U. & Müller-Hill, B. *EMBO J.* **2**, 593–597 (1983).
- Jacobson, R. H. & Matthews, B. W. *J. molec. Biol.* **223**, 1177–1182 (1992).
- Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489–501 (1987).
- Holmgren, A. & Brändén, C. I. *Nature* **342**, 248–251 (1989).
- Brändén, C. Q. *Rev. Biophys.* **13**, 317–330 (1980).
- Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P. & Dupourque, D. *Nature New Biol.* **238**, 37–41 (1972).
- Celada, F., Ullmann, A. & Monod, J. *Biochemistry* **13**, 5543–5547 (1974).
- Ring, M. & Huber, R. E. *Arch. Biochem. Biophys.* **283**, 342–350 (1990).
- Gebler, J. C., Aebersold, R. & Withers, S. G. *J. biol. Chem.* **267**, 11126–11130 (1992).
- Cupples, C. G., Miller, J. H. & Huber, R. E. *J. biol. Chem.* **265**, 5512–5518 (1990).
- Ullmann, A., Jacob, F. & Monod, J. *J. molec. Biol.* **24**, 339–343 (1967).
- Langley, K. E., Villarejo, M. R., Fowler, A. V., Zamenhof, P. J. & Zabin, I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1254–1257 (1975).
- Weinstock, G. M., Berman, M. L. & Silhavy, T. J. in *Gene Amplification and Analysis* (eds Papes, T. S., Rosenberg, M. & Chirikjian, J. G.) 27–64 (Elsevier, New York, 1983).
- Zabin, I. *Molec. cell. Biochem.* **49**, 87–96 (1982).
- Henderson, D. R., Friedman, S. B., Harris, J. D., Manning, W. B. & Zoccoli, M. A. *Clin. Chem.* **32**, 1637–1641 (1986).
- Adams, R. M. et al. *J. biol. Chem.* **269**, 5666–5672 (1994).
- Heinz, D. W. & Matthews, B. W. *Prot. Engng* **7**, 301–307 (1994).
- Sakabe, N. *Nucl. Inst. Meth.* **A303**, 448–463 (1991).
- Higashi, T. *J. appl. Cryst.* **22**, 9–18 (1989).
- Terwilliger, T. & Eisenberg, D. *Acta crystallogr.* **A39**, 813–817 (1983).
- Rayment, I. *Acta crystallogr.* **A39**, 102–116 (1983).
- Kraulis, P. J. *J. appl. Cryst.* **24**, 946–950 (1991).
- Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).
- Stokes, H. W., Betts, P. W. & Hall, B. G. *Molec. Biol. Evol.* **2**, 469–477 (1985).
- Buvinger, W. E. & Riley, M. J. *Bact.* **163**, 850–957 (1985).
- Schroeder, C. J., Robert, C., Lenzen, G., McKay, L. L. & Mercenier, A. J. *gen. Microbiol.* **137**, 369–380 (1991).
- Schmidt, B. F., Adams, R. M., Requadt, C., Power, S. & Mainzer, S. E. *J. Bact.* **171**, 625–635 (1989).
- Hancock, K. R. et al. *J. Bact.* **173**, 3084–3095 (1991).
- Gaykema, W. P. J. et al. *Nature* **309**, 23–29 (1984).
- Rossmann, M. G. et al. *Nature* **317**, 145–153 (1985).

ACKNOWLEDGEMENTS. We thank S. Clarke and M. Krevolin for cell fermentation; J. A. Wozniak and S. Snow for assistance with protein purification and crystallization; A. Nakagawa and N. Sakabe for their help at the Photon Factory; R. Albright, E. P. Baldwin, M. Blaber, S. L. Roderick, W. Tulip, L. H. Weaver and K. P. Wilson for help with data collection; S. L. Roderick, I. Rayment and M. G. Rossmann for technical advice; D. E. Tronrud for modifications to TNT; B. Branchaud and L. Blaszczak for synthesis of mercurated-c-galactose; L. Holm for database searches; and S. G. Withers, J. H. Miller, B. F. Schmidt and I. Zabin for discussion. This work was supported in part by a grant from the NIH to B.W.M.

CORRECTION

A map of a collisionally evolving dust disk around Fomalhaut

S. Alan Stern, Michel C. Festou & David A. Weintraub

Nature **368**, 312–314 (1994)

WE recently reported¹ mapping observations made in February 1993 indicating a detection of extended 1.3-mm emission around the nearby, main-sequence infrared excess star Fomalhaut. On 17 and 18 April 1994 we made similar but more sensitive observations of Fomalhaut at the same observatory (IRAM). These observations revealed that the extended emission we earlier reported appears to have been an artefact unique to the 1993 data. We therefore reprocessed the 1993 maps with improved facility software (R. Zylka, personal communication), and found that in each of the three 1993 maps there was a “source” near Fomalhaut, but that the precise position of this “source” changes by several arcsec from map to map. Thus, these sources cannot be real, but are probably due to variations in the thermal foreground to be of comparable size and brightness to a plausible source around the star. Apparently, the sum of the “sources” in the 1993 maps produced an extended object, fortuitously placed signal at the position of Fomalhaut on the sky.

On 18 and 19 April 1994 we made even more sensitive observations at IRAM using the same instrument in on-off (rather than mapping) mode; we used a chop-throw of 45 arcsec. These data reveal the detection of an unresolved, central source at a level of 7.1 ± 3.1 mJy, and a 3σ upper limit for a total flux of 51 mJy in the six, 20 arcsec-diameter channels that surround the central channel. This revised flux density contradicts our previous report¹ on the size, brightness, and physical characteristics of the Fomalhaut disk, and is consistent with results described in Scientific Correspondence on page 766 of this issue by Chini et al.² A detailed study of the ambiguities and past contradictions in millimetre-wave observations of very faint sources is reported in ref. 3. □

- Stern, S. A., Festou, M. C. & Weintraub, D. A. *Nature* **368**, 312–314 (1994).
- Chini, R., Krugel, E., Kreysa, E. & Zylka, R. *Nature* **369**, 766 (1994).
- Weintraub, D. A. & Stern, S. A. *Astr. J.* (in the press).